Socio-economic Inequalities in the Utilisation of Novel Anti-Cancer Therapies

Ruth P. Norris

A thesis submitted for the degree of Doctor of Philosophy

School of Pharmacy,
Population Health Sciences Institute

Newcastle University
February 2023

Abstract

Novel anti-cancer therapies use tumour biology to stratify patients on the likelihood of pharmacological response. Historically, conventional cancer treatment follows a socio-economic gradient. It is unknown if novel therapies are subject to similar inequalities. This thesis aims to determine whether there are socio-economic inequalities in the utilisation of novel anti-cancer therapies.

First, a systematic review and meta-analysis which synthesised evidence for socio-economic inequalities in novel therapy and predictive biomarker test utilisation was undertaken. Worldwide, low socio-economic status was associated with modestly low predictive biomarker test utilisation (OR 0.86, 95% CI 0.71, 1.05) and significantly lower precision medicine and biological therapy utilisation (OR 0.83, 95% CI 0.75, 0.91). Study data was primarily American, and inequalities varied by cancer type (larger in lung than breast cancers).

Second, an observational study was conducted using English population-based cancer registry data linked with the Systemic Anti-Cancer Therapy (SACT) dataset on 40,179 women, diagnosed with HER2+ breast cancer between 01/01/2012 - 31/12/2017. Multivariable logistic regression determined likelihood of trastuzumab utilisation by deprivation (measured by IMD quintile). For women living in the most deprived areas compared to the least deprived, modest trastuzumab utilisation inequalities were found (mvOR 0.92, 95% CI 0.85, 0.99).

Finally, a second observational study of the English population-based cancer registry data linked with the SACT dataset on 195,387 NSCLC cases diagnosed between 01/01/2012 - 31/12/2017 was performed. Multivariable logistic regression determined likelihood of any novel therapy utilisation by deprivation (measured by IMD quintile). For NSCLC patients living in the most deprived areas compared to the least deprived, significant novel therapy utilisation inequalities (mvOR 0.54, 95% CI 0.50, 0.58) were observed.

Despite treatment advances and regardless of healthcare system or cancer type, a low socio-economic status was associated with reduced novel anti-cancer therapy utilisation. Further work should explore why these inequalities occur to develop equitable approaches to therapy utilisation.

Dedication

Ada,

Thank you for slowing me down (in the nicest possible way). You have reminded me of the important things in life, as well as the reasons why I began the PhD in the first instance: to use time wisely, to grow and to enjoy what you do.

Keep being curious lovely,
Mum

Acknowledgments

The following list incorporates the people who have made the PhD years so memorable and who have not only shaped the ideas within this thesis, but also me.

I would like to express my gratitude to my supervisory team (Professor Adam Todd, Professor Linda Sharp, and Dr. Alastair Greystoke). Thank you for taking a chance on me, way back now it seems in the early Autumn of 2018, as the person to bring your ideas to fruition. I have taken a lot from your varied guidance, and I hope you are proud of our accomplishments. Adam, you've always championed my work and pushed me to aim high. Linda, your teaching, encouragement, and investment in me has made me all the better a researcher. Ali, I'll never forget the time I spent time with you in clinic. It was a great reminder that behind the information presented in this thesis, or the data the captured in the electronic healthcare records, are people. People with stories and lives. The human element of healthcare research. Thank you for sharing that with me. I am hopeful that this work marks some small step towards improving cancer care - for all.

I am grateful for the support of my internal reviewers, Professor Ruth Plummer, and Dr. Svetlanna Glinianaia. I enjoyed the short time we spent together in annual meetings where thoughtful suggestions and constructive review were provided. You both always took a great interest in me - as well as the thesis. Thank you is also extended to the Barbour Foundation for funding the PhD scholarship.

The research would not have been possible without the assistance of many colleagues and research partners. This included: Mark Bousfield, Henry Cain, Nicola Cresti, Antonia Hamill, Alex Inskip, Professor Kristina Johnell, Sarah Lawton, Dr. Agieszka Lemanska, Dr. Lucas Morin, Stephen Rice, Jonathon Smith, Dr. Katie Thomson, and all the patients who so kindly (and no doubt inadvertently) shared their pseudonymised data for this analysis in the hopes that cancer research can improve care for those living with cancer or who are yet to be diagnosed.

I have been fortunate to have worked alongside some excellent early career researchers and academics who provided mentorship despite otherwise traditional academic hierarchy. First, mention is made to Dr. Alice McCloskey for her guidance, friendship and reinforcing something that I knew all along when I signed up for the PhD and which the Irish fondly refers to as "making your dash count". You've been a great miss since you ventured West. The second thank you is extended to Dr. Rosie Dew for always being limitlessly generous with her time

and having patience with me, Stata, and all things statistics. Third, to Dr. Adam Rathbone – I'm extremely grateful that the PhD afforded us both the opportunity to become great friends. I have many a fond memory of the Northern Stage and Les Petits Choux days. Thank you for taking a real interest in my pastoral support.

I am also grateful to my fellow PhD peers in both the School of Pharmacy (Anna Robinson, Anum Iqbal, Emma Smith, Janine Beezer, Jawza Alotaibi, Lara Ahmaro, and Tania Jones) and the pre-pandemic Sir James Spence office (Jennifer Deane and Nick Walton) for sharing in the ups and downs of daily PhD life. An extended thank you also goes to the wider (past and present) Cancer Team (Dr. Adam Biran, Ben Rimmer, Dr. Christina Dobson, Dr. Fiona Menger, Dr. Lisa Crowe, Dr. Liya Lu, Dr. Lizzie Dutton, and Dr. Morven Brown) for welcoming me into the research family, providing me with a much needed second (statistical) home as well as social (Zoom) contact once the pandemic would hit. Additional mention is made to the School of Pharmacy staff for welcoming me back from my undergraduate days and for providing a base that I have always felt at home in. Finally, to the Ridley crew (Dr. Morgan Beeson, Cara Leavey, Dr. Charlotte Parbery-Clark, Tim Price, Amber Sacre, and Julija Simpson), thank you for your support during the final stages of this process.

My parents have long been my biggest supporters and this thesis marks the culmination of their academic investment in me. From an early age they have nurtured my curiosity, instilled a strong work ethic, and continued to highlight that the outputs from a quiet and continued determination should never be overlooked. Additionally, thank you to my in-laws (Paula and Neil) for the endless number of hours spent with Ada to enable me to write up.

Personal mention is expressed to my friendship group: Adam, Alice, Anna L, Catherine, Emma S (x2), Emily, Gillian, Helen, Isobel, Jenefer, Laura, Natalie, Nicola, and Steph. These are the people with whom I have shared both my professional and personal highs and lows over the course of the PhD. They are the people who were most resolute that I would "finish" and arguably also most looking forward to relishing in the accomplishment of having "Dr. Ruth" in the friendship group.

Finally, a lot of life has also happened during the formation of this body of work. My final thank you is therefore for my husband Mark - my unofficial fourth supervisor, sounding board, mentor, and best friend. Feeling incredibly fortunate for your unwavering support, belief in me and patience throughout the process.

Publications & Conference Presentations from the Thesis

Peer Reviewed Papers

Norris. R. P, Dew. R, Sharp. L, Greystoke. A, Rice. S, Johnell. K and Todd. A (2020) Are there socio-economic inequalities in utilisation of predictive biomarker tests and biological and precision therapies for cancer? A systematic review and meta-analysis, BMC Medicine 18, 282, https://doi.org/10.1186/s12916-020-01753-0.

Abstracts

Norris. R. P, Dew. R, Todd. A, Greystoke. A. and Sharp. L (2022) Socio-economic inequalities in NSCLC treatment during the era of tumour biomarker guided therapy: A population-based study, J Thorac Oncol, 17 (9): 55-56, https://doi.org/10.1016/j.jtho.2022.07.017.

Blog Posts

Norris. R. P. (2020) Can your education, income or even your job affect your chances of receiving newer cancer treatments? FUSE Open Science Blog, https://fuseopensciencebloG.blogspot.com/2020/10/can-your-education-income-or-even-your.html, accessed 26/7/2021.

Conference Presentations

Norris. R. P. *et al.* (March 2020) Are there socio-economic inequalities in utilisation of novel cancer therapies? A systematic review, oral presentation at the European Drug Utilisation Research Group Conference, Szeged, Hungary.

Norris. R. P. *et al.* (November 2020) Are there socio-economic inequalities in utilisation of predictive biomarker tests and biological and precision therapies for cancer? A systematic review and meta-analysis, oral presentation, and poster at the NCRI Virtual Showcase.

Norris. R. P. *et al*. (November 2021) Fair Treatment for All? Socio-economic inequalities in HER2+ breast cancer treatment utilisation, oral presentation, and poster as the NCRI Virtual Festival.

Norris. R. P *et al.* (August 2022) Socio-economic inequalities in NSCLC treatment during the era of tumour biomarker guided therapy: a population-based study, oral presentation at IASLC 2022 World Conference on Lung Cancer, Vienna, Austria.

Table of Contents

List of Tables

List of Figures

List of Figures Continued

List of Abbreviations

adjOR	Adjusted odds ratio
AIC	Akaike information criterion
ALK	Anaplastic lymphoma kinase
ALKi	Anaplastic lymphoma kinase inhibitor(s)
Anti-CTLA4	Anti-cytotoxic T-lymphocyte-associated protein 4
APC	Admitted patient care
ASCO	American Society of Clinical Oncology
BCLC	Barcelona Clinic Liver Cancer Stage
BIC	Bayesian information criterion
BMI	Body mass index
BNF	British National Formulary
BRAF V600	V-raf murine sarcoma viral oncogene homolog B1 V600
BRCA 1/2	BReast CAncer gene 1/2
CALR	Cancer activity level reporting database
CAS	Cancer Analysis System
CCM	Charlson Comorbidity Index
CCR	California Cancer Registry
CDF	Cancer Drugs Fund
CI/CIs	95% Confidence interval(s)
CIHI-DAD	Canadian Institute of Health Information Discharge Abstract Database
CoC	Commission on Cancer
COPD	Chronic obstructive pulmonary disease
CORECT-R	COloRECTal cancer repository
COSD	Cancer Outcomes and Services Dataset
CPRD GOLD	Clinical practice research datalink
CQC	Care Quality Commission
CRC	Colorectal cancer
CRN	Cancer Research Network
CRPC	Castrate-resistant prostate cancer
CRUK	Cancer Research UK
CTL1-4	Choline transporter-like protein 1-4
DCIS	Ductal carcinoma in situ
DCO	Death certificate only
DNA	Deoxyribonucleic acid
DOB	Date of birth
EC	Epirubicin and cyclophosphamide
ECOG	Eastern Cooperative Oncology Group Scale
EGFR/EGFRi	Epidermal growth factor receptor/(inhibitor)
EGFR-TK	Epidermal growth factor receptor tyrosine kinase
EHR(s)	Electronic health record(s)
EMA	European Medicines Agency
EML4-ALK	Echinoderm microtubule-associated protein-like 4 anaplastic lymphoma
ENCORE	English National Cancer Online Registration Environment
ER	Oestrogen receptor
ESMO	European Society for Medical Oncology
FISH	Fluorescence in situ hybridisation
FDA	Food and Drug Administration Federal Agency
FEC	Fluorouracil, epirubicin, and cyclophosphamide
FLOT	Fluorouracil, leucovorin, oxaliplatin, and docetaxel

GDP	Gross domestic product
HbA1c	Haemoglobin A1c
HER2/HER2+/-	Human epidermal growth factor receptor 2 (positive/negative)
HES	Hospital episode statistics
HES-APC	Hospital episode statistics admitted patient care
HMO	Health maintenance organisation
HSPC	Hormone-sensitive prostate cancer
ICD	International classification of disease
ICES	Institute for Clinical Evaluative Sciences
ICI/ICIs	Immune checkpoint inhibitor(s)
ICL	Inverse care law
ID	Identifier
IGI/IGIs	Intervention generated inequality/ies
IHC	Immunohistochemistry
IL-2	Interleukin-2
IMD	Index of multiple deprivation
IRAS	Integrated research application system
IRR	Incidence rate ratios
IRSD	Index of relative socio-economic disadvantage
ISPOR	International Society for Pharmacoeconomics and Outcomes Research
KIT	Feline sarcoma viral oncogene v-kit
KPIs	Key performance indicator(s)
KPS	Karnofsky Performance Status
KRAS	Kirsten rat sarcoma virus
KRAS G12c	Kirsten rat sarcoma virus glycine-to-cysteine substitution at codon 12
LRT	Likelihood ratio test
LSOA	Lower layer super output area
MAB/MABs	Monoclonal antibody/monoclonal antibodies
MBS	Medicare benefits scheme
MDT	Multidisciplinary team
MET	Mesenchymal epithelial transition factor receptor
MHRA	Medicine and Healthcare products Regulatory Agency
M-H	Mantel-Haenszel
MLH1	MutL homolog 1
MSH2	MutS homolog 2
MSH6	MutS homolog 6
mvOR	Multivariable odds ratio
NACRS	National ambulatory care reporting system
NBOCA	National bowel cancer audit
NCCN	National Comprehensive Cancer Network
NCDB	National Cancer Database
NCI	National Cancer Institute
NCRAS	National Cancer Registration and Analysis Service
NCRD	National Cancer Registry Database
NDFP	New drugs funding program
NDRS	National Disease Registration Service
NHS	National Health Service
NICE	National Institute for Health and Clinical Excellence
NLCA	National lung cancer audit
NOGCA	National Oesophago Gastric Cancer Audit
NOS	Not otherwise specified
NR	Not reported

NRAS	Neuroblastoma ras viral oncogene homolog
NSCLC	Non-small cell lung cancer
NTRK 1/2/3	Neurotrophic tropomyosin-receptor kinase 1/2/3
NYSCR	New York State Cancer Registry
OCISS	Ohio State Incidence Surveillance System
OCR	Ontario Cancer Registry
ODB	Ontario Drugs Benefit
ODR	Office for Data Release
OHIP	Ontario Health Insurance Plan
ONS	Office for National Statistics
OR/ORs	Odds ratio(s)
PAPs	Patient assistance programmes
PARP	Poly (ADP-Ribose) Polymerase
PAS	Patient administration systems
PBS	Pharmaceutical benefits scheme
PD-L1	Programmed cell death protein 1/ligand 1
PHE	Public Health England
PICOS	Population, Intervention, Comparison, Outcome and Setting
PI3K	Phosphatidylinositol 3-kinase
PMS2	PMS1 homolog 2, mismatch repair system component
POC	Patterns of care
POLD1	DNA polymerase delta 2 catalytic subunit
POLE	Polymerase epsilon catalytic subunit
PR	Progesterone receptor
PRISMA	Preferred Reporting Items for Systematic Reviews and Meta-Analyses
PROSPERO	International database of prospectively registered systematic reviews
PSM	Propensity score matching
QA	Quality appraisal
QOF	Quality and Outcomes Framework
R-CHOP	Rituximab, cyclophosphamide, doxorubicin, vincristine, and prednisolone
RCT	Randomised controlled trial
RET	Rearranged during transfection
REC	Research ethics committee
ROS-1/ROS1	Receptor tyrosine kinase 1
R-miniCHOP	Rituximab and reduced dose CHOP
RPDB	Registered persons database
RR	Risk ratio
RTSD	Radiotherapy dataset
SABR	Stereotactic ablative body radiotherapy
SACT	Systemic Anti-Cancer Therapy/ies (dataset)
SCLC	Small cell lung cancer
SCR	Summary care record
SD	Standard deviation
SEER	Surveillance, Epidemiology and End Results program
SEP	Socio-economic position
SES	Socio-economic status
SDoH	Social determinants of health
SMARC4	SWI/SNF related, matrix associated, actin dependent regulator of chromatin, subfamily A, member 4
SPC	Summary of product characteristics
TA	Technology appraisal NICE guidance number

TAC	Docetaxel, doxorubicin, and cyclophosphamide
TCH	Docetaxel, carboplatin, and trastuzumab
TKI/TKIs	Tyrosine kinase inhibitor(s)
TNBC	Triple negative breast cancer
TNM	Tumour, node, metastases
T790M	Gatekeeper mutation of the epidermal growth factor receptor
TX	Texas
USA	United States of America
VICORI	Virtual Cardio-Oncology Research Initiative
VEGF	Vascular endothelial growth factor
WHO	World Health Organisation
YCR BCIP	Yorkshire Cancer Research Bowel Cancer Improvement Programme
2WW	Two week wait

List of Definitions

Biological/Biologic Therapy	Specific type of targeted therapy which has no associated predictive biomarker status included in the licence. Taken to refer to anti-angiogenics or any broad literature reference to a "targeted biologic".
Biomarker Test	A test for a biological molecule found in blood, other bodily fluids, or tissues that is a sign of a normal or abnormal process, or of a condition of disease.
Conventional Treatment	Refers to surgery, chemotherapy and/or radiotherapy.
Health Inequality	Unfair and avoidable differences in health across the population, and between groups within society. Definition includes healthcare inequalities e.g. treatment access.
Immunotherapy	A treatment that helps the immune system recognise and attack cancer cells.
Socio-economic Status (SES)	A measure of an individual or group's combined economic and social standing.
Novel Anti-Cancer Therapy	Refers to precision medicines (molecular targeted therapies), biologicals (anti-angiogenics) and immunotherapies.
Molecular Targeted Therapy	A drug or other substances used to target molecules involved in the growth and spread of cancer cells.
Precision Medicine	Refers to customisable treatments which consider variability in genes, environmental and lifestyle factors shared by a sub-group of patients.

Chapter 1. Introduction

1.1 Introduction

In 1999, the Wall Street Journal published an article describing a "new era of personalised medicine" directed by tailored treatments based on individual genetic makeup (1). The promise of this treatment revolution was large (complete transformation in medicine) and the implications for patients, potentially huge (reduced toxicity, improved survival) (2). In the twenty years or so that have since passed, the field has moved on considerably. Targeted treatments such as trastuzumab, imatinib and gefitinib are now standard practice in cancer and beyond. Additionally, predictive biomarker testing has now become a routine part of clinical practice. Though the terminology has changed somewhat (precision medicine and stratified medicine are now preferred) (3,4), the concept of using pharmacogenomics to direct and improve cancer treatment is here to stay. The extent to which this change has resulted in equitable treatment utilisation for all patients, regardless of socio-economic background remains unclear. This thesis examines whether the previously noted socio-economic inequalities present in conventional cancer treatments persists in novel anti-cancer therapies. Specifically, it deals with determining whether these precision medicines are subject to socio-economic inequalities in their utilisation.

Chapter Aim: To provide an overview of the thesis research topic.

Objectives:

1. Describe the relevant background to the research topic.
2. Provide a rationale for the thesis.
3. Define the research question and explain its significance.
4. Define the scope of the thesis.

The Chapter starts with a general overview of health and healthcare inequalities. Cancer terminology is then introduced along with the rationale for focusing on breast and lung cancers. A discussion of inequalities in cancer with a focus on socio-economic associations with mortality and survival then follows. A summary of cancer treatments (conventional methods and novel anti-cancer therapies) is then provided. This is followed by a short discussion on whether precision medicine may provide a solution to addressing known inequalities in conventional cancer treatments - thus setting up the aim and scope of this work. An overview of the thesis then concludes the Chapter.

1.2 Health Inequalities

Health Inequalities are defined by the King's Fund as "avoidable, unfair and systematic differences in health between different groups of people" (5). Fundamentally, these health inequalities relate to differences in the status of peoples' health. However, the term is used interchangeably to refer to opportunities which contribute to health status and by this extension, a health inequality can also refer to differences in access to care – for example in the availability of drug treatments (5).

There are many kinds of health inequality (e.g. age, gender, race/ethnicity) and ways in which the term is used. This thesis explores treatment utilisation inequalities between groups with different socio-economic statuses.

Health inequalities exist across the life course (6,7), occur in all countries and over many measures of health (8), morbidity and mortality (9). These inequalities are found across the entirety of the social gradient – so even those grouped in the middle have poorer health than those above and better health than those below (10). Health inequalities are problematic when avoidable circumstances lead to disproportionate disadvantage for those most in need (11). Eradicating health inequalities and creating a fairer society is therefore not only a matter of social justice but a priority; everyone should be entitled to the right to lead a healthy life (12). However, improving population health has proved complex and health inequalities have been observed to persist even after concerted efforts to address them (e.g. universal free at the point of service healthcare) (12). The Nordic States provide a prime example, as despite generous welfare coverage and overall high life expectancy, paradoxically inequalities still exist (13,14).

1.2.1 Socio-economic Inequalities

Socio-economic status (SES) is defined as a measure of an individual or groups' combined economic and social standing (15). The term historically references social class as a hierarchical structure determined by individual, relative position accounting to control over wealth, prestige and power (16,17). Nowadays, SES can be considered to measure occupational class, income or education level (16) and infers capacity to access basic resources needed to achieve and maintain good health (18). SES is thus a characteristic by which health inequalities are observed (5).

There are numerous SES measures and debate over what each measure represents and how this may influence the inequalities observed (19,20). Example measures include the following:

income; wealth; poverty; deprivation; education; occupation; and composite indices. These can be classified as: (i) compositional (characterising the individual e.g. employment status); (ii) contextual (characterising the environment e.g. census tract zip code); and (iii) composite (several measures combined to form indices e.g. Index of Multiple Deprivation (IMD)) (18). SES is complex, can vary over the life course and be measured at different levels (individual, household and neighbourhood) (21). The terms socio-economic position (SEP), social stratification and social class are often used interchangeably with SES (19).

There is no single or best choice of indicator (21). However, choosing the best measurement of SES for the study in question is important. Scholars need to consider both data availability and how the SES measure used may influence the health of the population explored (22). It can therefore be useful to investigate the advantages and disadvantages of individual SES measures before their application, especially within the wider context of likely causal pathways and outcomes for each study at that point in time (19,21,22). Table 1.1 lists such comparisons for three conventional single measures of SES (income, education, and occupation). Following such a comparison, though income is often considered the most useful measure of material goods and services protective to health, its application is less appropriate if sensitive individual data collection is restricted (21). Furthermore, ease of educational measures should not overshadow the measure's suitability if changes in educational opportunities (e.g. graduate job market post qualification) no longer provide a useful measure in that birth cohort of future earning and occupational potential (23). Finally, if it is not possible to assign occupational categorisation to the population in question (e.g. retirees), the strengths of this measure to determine social standing are eroded (22). Therefore, as all individual measures have limitations, it may be preferable to instead use multiple SES measures or a composite index which already combines different aspects of SES measurement (e.g. Carstairs Index) to provide the context that SES measurement is a complex construct.

Table 1.1 Advantages and disadvantages of conventional SES measures

	Advantages	Disadvantages
Income	• Useful single indicator of material goods and living standards which may influence health (22)	• Sensitive nature of information can result in reluctance to disclose at an individual level (24) • Often does not measure disposable income (19) • Age dependent (22) • Often does not also include health insurance, asset ownership (wealth) or disability benefits (22) • Not a proxy for wealth (21)
Education	• Easy to measure, often with a high response rate (25) • Relevant regardless of age or working circumstances (25) • Measurement captures aspects of lifestyle and behaviour (22) • Likelihood of reverse causation (if poor health or low SES came first) is reduced (22)	• Meaning varies by birth cohort due to changes in educational opportunities (19) • Less relevant if education gained in country outside of where currently reside (19) • Numerically quantifying education by years provides no information on experience (19) • SES does not necessarily consistently rise with years of education (22)
Occupation	• Data availability varies by data source (19) • Can provide measure of working conditions/psychological working demands (22)	• Classification issues if not currently employed or self-employed (19) • May not reflect current social circumstances e.g. retirees & homemakers (22) • Tends to underestimate SES variation if used as a single measure (19) • Some occupations can be unwilling to disclose (19) • Measurement lacks precision (22) • Lacks consideration of racial and gender occupational variation (22)

Socio-economic inequalities affecting health refer to avoidable, unequal and unfair differences in health status between different socio-economic groups (26). These health differences can be observed as: a lower life expectancy, higher overall mortality, and higher rates of infant and perinatal mortality for low socio-economic groups when compared to those in higher socio-economic groups, to give three examples (27). Factors associated with a higher socio-economic status (e.g. occupational prestige, higher income and educated to degree level or higher) are associated with better health outcomes (12).

1.3 The Determinants of Health

Good health is not simply the result of genetics, individual behaviour and medical care - though these factors are important (28). Rather, there are several determinants which combine to affect health, and systemic variations in these determinants are important for generating health inequalities.

The determinants of health (many of which an individual is unlikely to have direct control over) include: (i) the social and economic environment; (ii) the physical environment; and (iii) a person's individual characteristics and behaviours (29). Combined these factors determine the ability of the individual to have physical, social, and personal resources to identify and achieve goals, meet their needs and deal with changes in circumstance.

1.3.1 Dahlgren-Whitehead Rainbow

The Dahlgren-Whitehead Rainbow Model (Figure 1.1) was proposed in 1993 to summarise the main determinants of health as rainbow layers of influence (30). At the model's core are individual compositional factors which are largely fixed (e.g. age and sex) (30). The layer surrounding these factors reflect personal behaviours which either promote or damage health and are known to be modifiable (30). Constitutional and lifestyle factors may be linked. For example, ethnicity/race may influence cultural diet choices and risk taking behaviours such as drinking alcohol (31). The third layer considers the role of social networks - both family-based and from the wider community. For example, high levels of social capital (being well connected and having networks with people of influence) has been shown to be protective of good health (32). Conversely, in areas where discrimination and segregation exist (e.g. based on race) social mechanisms such as education, employment, housing and other opportunities may result in poorer health (33). The fourth layer considers further mediators on health, including living and working conditions, job opportunities, food services, education and housing provision (34). The final layer details general contextual socio-economic, cultural and environmental factors which may be health promoting (e.g. green space, lack of air pollution, good infrastructure, disposable income) or not (28). All layers can interact and exposure to the determinants varies with SES. For example, gene function can alter throughout the life course on account of individual behaviour and environmental exposure - this is known as epigenetics. SES may also influence lifestyle exposures. For example, certain diets and smoking (which are socio-economically patterned) can promote deoxyribonucleic acid (DNA) methylation, which in turn may alter gene expression and increase susceptibility to disease (35,36).

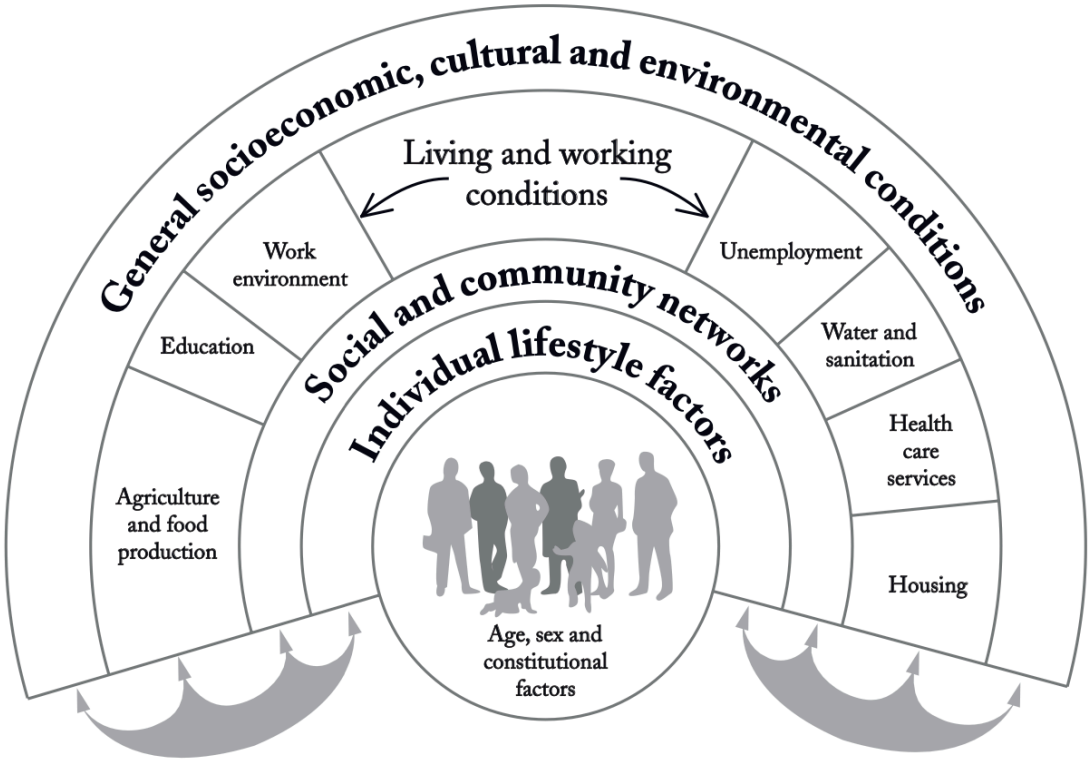


Figure 1.1 The main determinants of health. A conceptual framework devised to outline population health determinants. Taken from Dahlgren & Whitehead (1993) (30).

1.3.2 Social Determinants of Health

The idea that the conditions in which people are born, live, work and age influence health are referred to as the Social Determinants of Health (SDoH). These broad social and economic factors which determine health were also documented in the Dahlgren and Whitehead's Rainbow Model. The SDoH refer to the "causes of causes" of ill health and inequality. The main SDoH are: (i) working conditions; (ii) unemployment and worklessness; (iii) access to goods and services (water, food, and sanitation); (iv) healthcare access; and (v) housing. Variation in an individual's social class leads to differential exposure to the SDoH.

The SDoH also feature in "Fundamental Cause Theory", developed by Link and Phelan in 1995 to explain why socio-economic gradients reproduce over time, despite changes in factors that are thought to influence disease (e.g. the introduction of new medical intervention) (27). The theory proposes that resources such as money, knowledge, prestige, power, and beneficial social conditions are the fundamental drivers of disease (27). Inequalities therefore persist as medical advances do little to alter these structural determinants of disease and over time, the driving mechanisms find new means by which to further perpetuate the health gap (37). In essence, implementation of "new" interventions downstream in the system fails to account for the

unequal distribution of income, wealth, and power and the SDoH (fundamental causes) upstream (27).

1.3.3 Other Socio-economic Determinants of Health

Several other theories (38) build on the SDoH and seek to explain specifically why socio-economic inequalities may arise. These are described below.

Materialist Theory: This theory argues that a good income provides access to health benefiting goods and services (e.g. healthcare, schools, transport, social care) and limits exposure to material risk (e.g. poor housing, inadequate diet, physical work hazards and environmental exposure) which can harm health (31). Under such theory, patients living on low incomes may struggle to: (i) meet basic needs (obtain food, shelter and medicine); (ii) may live in less desirable and dangerous communities (e.g. substandard housing, high crime rates, lack of amenities); and (iii) have to endure more daily hardships of life that unaffordable, material items would otherwise ameliorate. Over time, these conditions serve to undermine any individual health promoting behaviours (39). A lack of material resources can thus increase an individual's "allostatic load" which in turn, results in chronic stress and further exacerbates the physical "wear and tear" of good health (40).

Behavioural-cultural Theory: This theory posits that health damaging behaviours (e.g. poor diet, drinking alcohol, smoking, lack of physical activity) are higher among lower socio-economic groups and that unhealthy behaviour is more culturally acceptable (26,31).

Psychosocial Theory: Psychosocial theory links health to the unequal social distribution of psychosocial risk arising from levels of control at work (e.g. repetitive tasks, poor job security, unemployment) and support in the community or in social spheres (e.g. family conflict, childhood abuse, divorce) (31). Combined these factors impact emotional development, psychiatric health and risk taking behaviour (26). Such a theory highlights the importance of mental health on physical health (41). The theory helps explain why a high SES has been found to correlate with improved psychological coping (42) or why low income individuals are at increased risk of experiencing emotional stress accounting to rising healthcare and insurance costs coupled with a reduction in employment and income worries post a disease diagnosis (43).

Life Course Theory: This framework considers the accumulation of disadvantage (social, psychological and biological) over the life course as opposed to viewing different aetiological

factors that an individual experiences at different times (38). Under this theory, inequalities arising early in life can predispose individuals to further inequality later on (26). Additionally, these exposures may be multiple or vary in magnitude and composition given engrained SES gradients (44). For example, a lower SES experienced during childhood has been shown to predict educational attainment and adult mortality (42).

The literature shows relatively few attempts to judge the contribution of each theory to observed health inequality generation, and those studies which do exist have failed to reach consensus (45). For example, a Norwegian cohort study exploring the role of material, psychosocial, behavioural, and biomedical factors in mortality concluded that material factors were most important in accounting for male income inequalities, whereas psychosocial and behaviour factors were more relevant to educational inequalities (46). A similar finding regarding the relevance of material factors was also found in a prospective observational study from the Netherlands assessing the direct and indirect contributions of material, behavioural, and psychosocial factors of education inequalities in mortality (47). In contrast, a Dutch cross-sectional study considering material, behavioural, cultural, and psychosocial factors in adult oral health concluded that behavioural contributions were most relevant, even above any material factors. Cultural and psychosocial mechanisms of health inequality were found to be moderate (48).

In contrast, there is a larger body of literature judging the influence of a single theory as the root cause of health inequality. By volume, behaviouralist theory, with its substantial study numbers could therefore be considered strongly supported especially when compared to other inequality theories (49,50). However, as the theory is also highly critiqued for its inability to address underlying, upstream causes of unhealthy behaviour development (e.g. education), the quantity of evidence is no substitute for quality when establishing the relative importance of behavioural theory (51). Cultural theory has similarly been critiqued as a poor fundamental explanation for health inequality generation. The evidence base for cultural mechanisms is thus considered weaker, even if its wider discussion does provide further insight into the complexity of health inequality generation (51). There is some argument that materialist/structural theory provides a more dominant, overall reasoning (51). However, materialist theory is also critiqued (inequalities are still observed in rich countries with material resources) and often compared to psychosocial theory, with scholars arguing that the perception of relative social hierarchy is just as crucial as exposure to any material advantage (47). Arguably, attempting to explain health inequalities as the result of a single theory is not useful (38). Rather, in order to avoid a

reductionist stance, it is more helpful to consider health inequality theories as overlapping (31) given that exposure to one factor (e.g. material) may result in exposure to another (e.g. psychosocial) (31,45). In this sense, the evidence for each theory should be considered within the wider context of all other mechanisms of inequality generation in that instance. Arguably, life course theory aims to achieve such a standpoint; it compliments other theories (e.g. unhealthy behaviour is important) but also emphasises the relevance of a temporal context (e.g. duration of unhealthy behaviour across the life span) (44).

1.3.4 Healthcare Service Provision & Access

In addition to SES and the SDoH, healthcare access and services provision (both disease prevention and treatment) has a role in explaining inequalities. Examples of these healthcare factors include: treatment access; funding (social insurance, private or general taxation) (31); academic versus community healthcare (52); hospital case volumes; availability of services (hours of operation, referral systems, rural versus urban settings); guidelines; participation in clinical trials (53); and discussion at multidisciplinary team (MDT). Healthcare system biases have also been documented (age, gender and race) and this may further limit access to services and treatments (18). However, good health is not simply a question of good healthcare service or its provision. For example, in the USA (an affluent nation), still observes lower life expectancy and higher mortality despite significant expenditure on healthcare (54).

The Inverse Care Law (ICL), as first defined by Hart in 1971 (55), describes healthcare access. The law proposes that the availability of good medical care varies inversely with the needs of the population served (Figure 1.2). This law is often seen as synonymous with SES variation given the close association of health and healthcare needs (56). Under the law, those in low socio-economic groups are less likely to access and use healthcare services even when there is no difference in the level of health need to high socio-economic groups (55). The ICL has been found to exist in a range of circumstances and highlights the disadvantage on lower socio-economic groups (57).

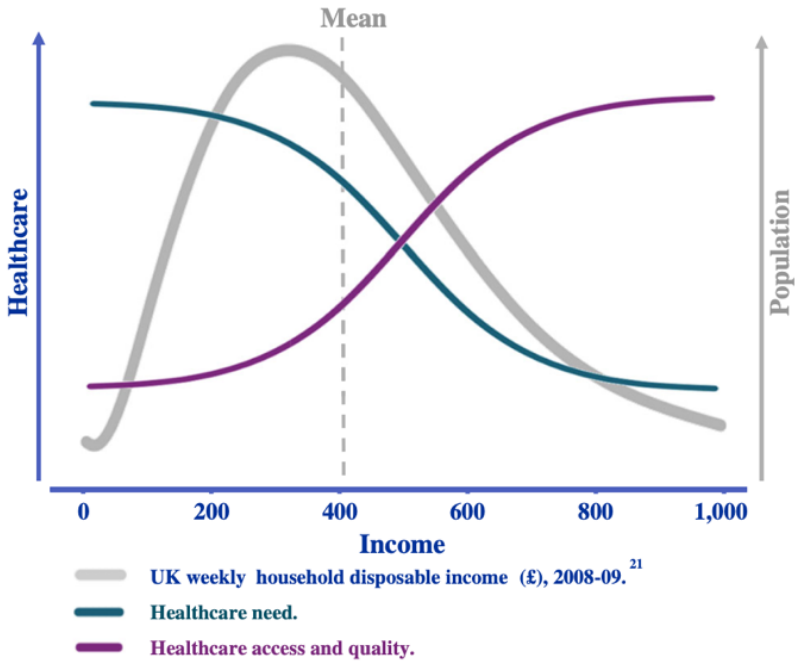


Figure 1.2 Graphical depiction of the ICL. Income distribution for the population is skewed towards lower levels, yet healthcare need varies inversely with income and the lowest incomes have higher healthcare need. Taken from James *et al.* (2014) (57).

1.3.5 Complex Determinants of Health

Combined the determinants listed indicate the many influences on health - biology, physical and social environment, personal life, and health services. There is no consensus on which factors are most responsible for determining good health (58). It can therefore be helpful to take an "intersectionality" approach - that it, to consider the multiple influences on health and how these variables interact to form axes of inequality (31). It may also be useful to consider poor health as additive, if not multiplicative of individual risks (59).

SES is thus an important health and treatment determinant, which acts through a myriad of pathways (26). The multifaceted nature of such social causation in inequality is referenced in the term, "fundamental cause". This means that the effects of SES are amplified by risk factors associated with low SES clustering within families and communities as they likely share the same status and reside in the same low SES community (26).

1.4 Cancer

Globally cancer is a major public health concern (60). One in two people are predicted to develop cancer in their lifetime and whilst treatment has improved outcomes, prognosis for many, remains poor (61,62). Data from 2020 reported an estimated 19.3 million new cancer cases and 10 million cancer deaths worldwide (60). Incidence is projected to rise to an estimated

28.4 million by 2040, as a result of ageing populations and the high (and, in some instances, increasing) prevalence of cancer risk factors (e.g. smoking, alcohol, obesity) (60).

Cancer is a generic term for a large, complex group of heterogenous diseases characterised by uncontrolled growth and spread (metastases) of abnormal cells (63). Cancers are classified by the originating tissue (histology) and body location (primary site). The International Classification of Diseases (ICD) provides a naming standardisation for tumour site (topography) and histology (morphology) which is commonly used by cancer registries (64). Histologically, cancers can be grouped as: carcinoma (solid tumours with an internal or external lining of the body origin); sarcoma (connective tissue cancer); leukaemia (bone marrow cancer); lymphoma and myeloma (immune system cancers); and mixed type (65). As the definition is wide, this thesis will only focus on solid tumours, with later Chapters exploring two examples of these: breast and lung cancers respectively.

Cancer is a genetic disease and tumours are the result of mutated genes developing from combination of: (i) normal cell cycle division errors; (ii) carcinogen (e.g. tobacco) induced DNA damage; and (iii) genetic predisposition (66). Cancers are modulated by an array of: genetic; molecular; cellular; tissue; population; environmental; and socio-economic factors that evolve with time (62). Important cancer genetic mutation drivers include proto-oncogene activation (instructs cells to grow uncontrollably) and tumour suppressor gene inactivation (disrupts processing which ordinarily slows cell division) (67).

Novel anti-cancer therapies have been applied to many cancers, however breast (specifically, human epidermal growth factor receptor positive (HER2+) subtype) and lung (specifically, non-small cell lung cancer (NSCLC)) cancers provide two illustrative and contrasting contexts for investigation. These cancers form the basis for much of the work undertaken for this thesis and are referred to throughout the remainder of this Chapter. The rationale for choosing these cancers is due to their differing novel anti-cancer therapy applications, which provides for two illustrative cases studies.

Breast cancer, specifically the HER2+ subtype, documents one of the first successful and well-known examples of the application of a novel anti-cancer therapy in cancer - trastuzumab. This monoclonal antibody (MAB) became the first, blockbuster novel anti-cancer therapy approved (licensed by the Food and Drug Administration Federal Agency (FDA) & European Medicines Agency (EMA) in 1998 and integrated into the National Institute for Health and Clinical

Excellence Guidance (NICE) guidance in 2002, initially for advanced disease) (68). The development of trastuzumab informed the evolution of further targeted drug discovery (69). Today, trastuzumab remains on the World Health Organisation's (WHO) designated list of "essential medicines" (70). A focus on human epidermal growth factor receptor positive (HER2+) breast cancer thus provides a case study into a cancer subtype where a key therapy has marked somewhat of a dawn in a new treatment era and where today novel anti-cancer therapy practice is now engrained in clinical settings owing to its early licensing.

In lung cancer, adoption of novel anti-cancer therapies has been just as transformative, with patients benefiting from the substantial increases in targeted treatments and biomarkers tests that have helped distinguish the management of different histological sub-types (71,72). Introduction of tyrosine kinase inhibitors (TKIs) to target epidermal growth factor receptor (EGFR) mutations marked the start of this new treatment era back in 2004 (73). Since then, several novel anti-cancer therapies have been approved across a range of driver mutations. A focus on NSCLC thus allows for consideration of a range of novel anti-cancer therapies across several tumour mutations. Additionally, as much of the progress in novel anti-cancer therapy application in NSCLC has been in the stage IV population, its study also provides examination of utilisation in what has historically been a population group with limited treatment options.

1.4.1 Breast Cancer

Detection & Diagnosis: Breast cancer is generally diagnosed following either mammogram screening or a symptom detection (such as pain or palpable mass) that prompts further examination and referral (74). Population-based screening programmes were introduced in the 1980s and are now well established in Western countries (usually offered screening from the age of 40-50 upwards to age 70 on a one, two or three year basis) (75). In England, between 2020-2021, 10,813 women had breast cancers detected through screening (a rate of 9.1 cases per 1,000 women screened) (76).

Sub-types: Breast cancer, whilst traditionally considered to be a single disease, is actually characterised by several, distinct molecular sub-types based on hormone receptors (oestrogen (ER) and progesterone (PR)) and HER2 status (77). Sub-types are as follows: Luminal A (ER+ and/or PR+, HER2-); Luminal B (ER+ and/or PR+, HER2+); and Basal Cell Like (ER-, PR-, HER2-) which is also known as Triple Negative Breast Cancer (TNBC) (78). Chapter 3 of this thesis focuses on HER2+ breast cancer. HER2 (also known as Neu, ErbB2) is a member of the EGFR tyrosine kinase family which regulates cell differentiation and proliferation (79).

Overexpression of HER2 results in an aggressive breast tumour type which is thought to account for approximately 20-25% of all breast diagnoses worldwide (80).

Incidence, Mortality & Survival: In 2020, breast cancer reached a new milestone - exceeding the number of global lung cancer cases for the first time (81). Breast cancer is now the most commonly diagnosed cancer worldwide (an estimated 2.3 million new cases, representing 11.7% of all worldwide cancer cases in 2020) (60). Breast cancer can occur in men, though this is rarer (affecting approximately 100 times fewer men than women) (82) and subsequent comments in this thesis relate to breast cancer in women. Prognosis following timely diagnosis is generally good. Countries such as the UK and USA with established and effective treatment practices have thus observed a decline in mortality since the late 1980s/early 1990s with organised screening in the UK having a role in this decline (81,83). Survival varies by stage at diagnosis; generally breast cancer prognosis is good (around 85% of women will survive their cancer for 5 years or more after diagnosis for all stages in England) (84), though metastatic carcinomas are generally incurable (85) (only 66% of stage IV breast cancer patients between 2013-2017 in England survived their disease by one year) (84). Hence worldwide, breast cancer remains the 5th leading cause of cancer mortality (60). In the UK, between 2016-2018 there were on average 55,920 new cases of breast cancers a year, and 11,547 deaths during this time period (86).

SES Associations: Socio-economic gradients in breast cancer are observed; women with a higher SES have higher risk of developing cancer (87). This association likely reflects the increased exposure of women with higher SES to breast cancer risk factors. Primarily, this pertains to reproductive factors accounting from lifestyle choices. Women of a higher SES have been shown to delay childbirth, have fewer children and use exogenous hormones (e.g. contraceptives, hormone replacement therapy) – all factors which increase the risk of oestrogen exposure (the primary breast cancer driver) (78,88). Additionally, women of higher SES can be considered to have a higher health literacy (educational experiences over time) (89), the consequence of which may be that such women attend breast cancer screening and/or more often recognise potential symptoms and access the healthcare system. Such action may increase the likelihood of a breast cancer diagnosis outcome (90). Socio-economic breast cancer mortality associations have proved contradictory over time. Historically mortality was lower in those women of low SES, but evidence now suggests that this may be other way around and that women of low SES experience higher mortality perhaps due to barriers in treatment access

(37,91,92). However, after diagnosis, poorer survival has long been a feature of women with a low SES (93).

Risk Factors: Incidence of breast cancer is associated with a number of risk factors and these include: advanced maternal ages and lower parity; menopausal hormone replacement therapy use; along with high alcohol consumption; obesity; physical inactivity; genetics; smoking; and hormonal contraceptive use (though this is subject to debate) (60,78,94,95).

1.4.2 Lung Cancer

Detection & Diagnosis: Early lung cancer is largely asymptomatic so detection can be hard in the absence of physical changes and patients often live with lung cancer for numerous years before diagnosis is made (96). Symptoms, when present, include: a cough; chest and shoulder pain; haemoptysis; dyspnoea; weight loss; hoarseness; and fever (96).

Sub-types: Histologically, lung cancer can be divided into NSCLC (approximately 85% of patients) and small cell lung cancer (SCLC) (approximately 15%) (97,98). NSCLC can be further classified as: adenocarcinoma (which compromises approximately 40% of cases); squamous cell carcinoma (25-30%); and large cell carcinoma (5-10%) (98). Adenocarcinoma and large cell carcinoma fall under a "non-squamous" classification. Chapter 4 of this thesis explores these sub-types in greater detail.

Incidence, Mortality & Survival: Lung cancer is now the second most diagnosed cancer worldwide (11.4% of cases in 2020) (60). Between 2016-2018 there were 48,549 new cases of lung cancer in the UK and 35,137 deaths annually (99). Patients often present at a late stage and consequently survival post 5 years is still poor (100,101) (only 16.2% diagnosed in England between 2013-2017 survived their disease by 5 years or more) (99). In men, lung cancer remains the leading cause of cancer mortality; in women it is second to breast cancer (60).

SES Associations: SES is associated with lung cancer, with a low SES being associated with an increased risk of: occurrence (102), mortality (103) and survival (104).

Risk Factors: Environmental and lifestyle factors are important for lung cancer aetiology. Smoking (including second hand smoke exposure) is the major risk factor (105) accounting for an estimated 72% of these cancers in the UK (99) - and one that is known to be associated with a low SES (106). Smoking prevalence is higher in people living in poverty, on low incomes and

with less than a high school education (106). Quitting smoking is also less likely when education status is low (106). Other risk factors include: genetics (106); the environment (e.g. burned heating fuel pollutants and poor air quality); occupational carcinogen exposure (e.g. asbestos, arsenic and chromium dusts, radon exposure); and infections (e.g. tuberculosis) (60,98). Lung cancer can develop in non-smokers, but the patient population varies demographically. Non-smoking lung cancers tend to occur in patients who are: younger, female, of East Asian ethnicity and with an adenocarcinoma histology (106–109). Emerging evidence suggests lung cancer incidence rates is higher in younger women compared to younger men though this is not explained by smoking sex differences alone (historically male incidence is higher owing to adopting smoking earlier and at higher rates) (110).

1.5 Healthcare Inequalities in England

The impact of health inequalities in England continues to be observed, despite the existence of a post war welfare state with a National Health Service (NHS) which was built on the founding principle that care is to be readily available and free for all. An example of such health inequalities is shown by differences in life expectancy. For example, when comparing the life expectancy (2009-2013) of a female from birth, born in Stockton-on-Tees (an area of high deprivation and low SES) is 74.6 years, whilst ten miles away in Yarm (an area of low deprivation and high SES), this is almost 10 years higher (84.2 years) (111).

It was during the 1970s that inequality concerns were raised and the Black Report was commissioned to investigate the cause(s) (34,112). The Black Report confirmed the presence of inequalities, highlighting that the gap between high and low social classes was widening and this had yet to be redressed by health or social services (113). The report proposed four possible explanations (artefact, natural or social selection, materialist/structural and cultural/behavioural) and concluded that structural considerations were key (113). These findings built on previous work such as the Whitehall studies, the first of which was a study of 17,530 civil servants (data from 1967-1977) that reported a steep, inverse relationship between employment grade mortality from coronary heart disease - along with the fact that these inequalities were not confined to differences between just the rich and poor (114). The second Whitehall study, with a new cohort of 10,314 civil servants (data from 1985-1988), would later replicate these findings despite the 20 years which had elapsed from the first study (115).

In 1998, publication of the Acheson Report, exploring the widening in health between those at the top and bottom of the social scale, concluded that the evidence gathered for the report

supported socio-economic explanations that stem from basic structures of English society were causal for health inequality generation (116). The 1997-2010 health inequalities strategy which emerged focused on four themes: supporting families; engaging communities in tackling deprivation; improving prevention and care; and tackling the underlying determinants of health (117). By 2010, the change in government saw the end of this health strategy, which despite some partial successes, was concluded to have failed to reach its own targets (a 10% reduction in inequalities in life expectancy and infant mortality) (112,117). Since then, further reports ("Marmot Review 2010: Fair Society, Healthy Lives'; "Health People, Healthy Lives, 2011"; "Health Equity in England: The Marmot Review 10 Years On 2020") have emerged. The Marmot Reports served to evaluate the lessons from the WHO determinants of health and the 2011 strategy paper aimed to tackle policy failures in public health by placing an emphasis on the role of the individuals' health choices (12,118,119). The Marmot Reports concluded that: (i) life expectancy has stalled; (ii) years lived in ill health have increased; and (iii) inequalities have widened (118). Austerity and reduced public spending are outlined as major contributors to these poor outcomes. Despite minimal success, efforts to rectify such social injustice remain a priority to date (see the NHS Long Term Plan) (120).

1.6 Cancer Inequalities

An area of English medical care that has long been subject to health inequalities is cancer. People residing in more deprived areas are more likely to be diagnosed with cancer (20,000 more cancer cases a year when compared to the least deprived areas) and have poorer survival (121). These engrained inequalities between low and high SES highlight the important context in which new precision medicine interventions are being applied - settings, where historically, care is in inequitable and determined by the conditions in which patients have grown, work(ed) and live(d) (122).

1.6.1 Socio-economic Inequalities in Cancer

Socio-economic inequalities affect health at all stages of the cancer continuum, from diagnosis through to end of life care (123). Individuals living in the most socio-economically deprived areas compared to those in the least deprived may have: different perceptions of cancer lifestyle risks (e.g. diet and alcohol) (122); experience reduced access to cancer prevention and screening services (124); and are more likely to be diagnosed with advanced disease (122). Additionally, low SES is associated with: non-receipt of conventional anti-cancer treatments (e.g. surgery); reduced likelihood of referral for early stage clinical trials (122); and poorer overall outcomes (e.g. increased emergency admissions in their final year of life) (125,126). The disadvantage

associated with a low SES is seen both within and between countries (127), as well as occurring across a range of cancers, including both breast (128,129) and lung cancer (130,131).

1.6.2 Socio-economic Inequalities in Cancer Treatment

SES impacts cancer treatment throughout the cancer pathway. Figure 1.3 has been adapted using the Dahlgren-Whitehead Rainbow (30) to illustrate the main SES-related determinants of cancer treatment from the point of healthcare system entry. At the centre of the model is treatment utilisation. The rainbow layers which surround this reflect the sequential steps (or potential barriers) that a patient must progress through from referral and/or screening through to successful cancer treatment receipt. The outermost layers depict SES and the eight contextual SES factors (health literacy, financial resources, occupation, lifestyle, overall health, support, geography, and decision making). Health literacy refers to: an individual's ability to access and use information, navigate the system, and to interpret clinical judgment. Financial resources refer to an individual's ability to: pay for treatment(s), take time off work to attend appointments and undergo treatment, and access (where appropriate) healthcare insurance. Occupation, when considered beyond its financial implications, refers to work related carcinogen exposure which could drive tumour mutations and subsequent decision-making regarding treatments based on tumour biomarker presence/absence. Lifestyle refers to: smoking history, alcohol intake, diet, exercise levels, and risk-taking attitudes. Overall health considers any other medical conditions that an individual may have which may predispose them to cancer or exclude them from treatment (side effects, medication interactions etc). Support relates to the following: social networks, psychological coping mechanisms, and levels of stress. Geographical considerations include where a patient lives and ease of access and travel times to (specialist) healthcare services. Decision making relates to the following: capacity, beliefs (including cultural), and previous experiences. These factors are referred throughout the remaining rainbow layers with the use of colour coding to highlight where each SES factor may have an impact on treatment receipt. Some factors are relevant to all steps e.g. health literacy. This is because an ability to understand and interpret the health system and medical terminology is required from the point of entry, through to screening and biomarker test result interpretation and up to establishing the risk/benefit of treatments. In contrast, other contextual SES factors are more relevant at certain steps within the process. For example, support is particularly pertinent at referral, diagnosis, and treatment receipt. Please note, there is an array of additional factors (e.g. clinician biases, cancer waiting times) which also influence treatment utilisation. Such factors have been omitted in Figure 1.3 as patterning with SES was not considered feasible.

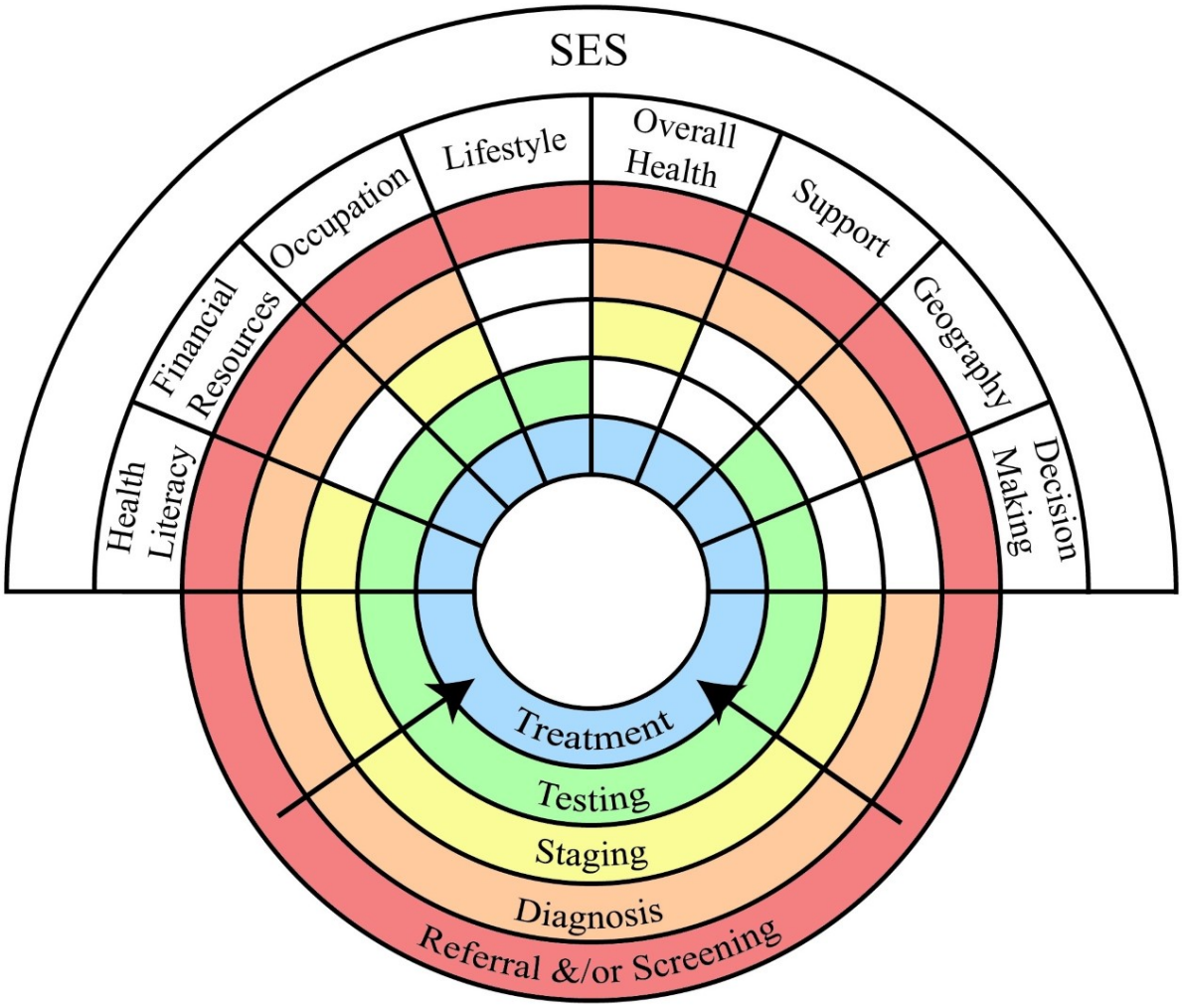


Figure 1.3 The main SES-related determinants of cancer treatment from the point of healthcare system entry. The shaded sections in the upper half of the figure indicate that the determinant may influence that aspect of the clinical pathway.

1.6.3 Socio-economic Inequalities in Cancer Mortality

For decades, cancer mortality in high income countries has been declining (91). However, this trend has not been experienced equally as those with a lower SES have witnessed increased cancer mortality and slower improvements in mortality rates (132). In England, cancer remains a major cause of mortality (133) and contributes to the deprivation gap between low and high SES (132).

In breast cancer, the increased screening programme and health service access has resulted in decreasing mortality (134). However, the association of SES with mortality has proved contradictory (88,135,136). Traditionally, mortality rates were lower in patients of lower SES (92,135). However, evidence is now emerging to indicate that mortality may now in fact be higher in women of low SES relative to those of high SES (91,137). In some settings,

differential treatment receipt has been proposed as an explanation (92). In England, there is evidence for a small association between female breast cancer mortality and deprivation in the period 2007-2011 with mortality rates being 6% higher for females living in the most deprived areas compared with the least deprived areas (138). Furthermore, the estimated deprivation gradient has not changed between 2002-2011 (138).

In lung cancer, greater mortality inequalities for patients with a lower SES compared to those with a higher SES have been observed (103,137,139,140). These inequalities have been shown to be more striking in the male population. Male lung cancer mortality has been in decline since the late 1980s but female mortality has not (141). In England between 2007-2011, in males and females, mortality rates were 170% higher for males living in the most deprived areas compared with the least deprived and 176% higher for women (138). The estimated deprivation gradient in lung cancer mortality in the most and least deprived areas has not changed between 2002-2011 (138).

1.6.4 Socio-economic Inequalities in Cancer Survival

A low SES has been associated with a reduced cancer survival time. This finding has been replicated in several clinical and geographical settings since the 1970s (142–144). In England, residence in a deprived area is associated with a lower cancer survival (145). The outcome of this has been avoidable excess deaths had survival in the most deprived groups been as high as in the least deprived (146). Reasoning for survival differences by SES are often summarised as a combination of: tumour (e.g. aggressive disease); healthcare (e.g. access to services); and patient factors (e.g. comorbidities) (143,146). Cancer stage is often stated as a key driving mechanism, though differential treatment between social groups has also been raised a major contributor (143). In England, a number of initiatives (e.g. NHS Cancer Plan and 2007 Cancer Reform Strategy) have tried to address socio-economic survival concerns related to a low SES, but despite their concerns, the deprivation gap has persisted (147).

In breast cancer, whilst there have been overall improvements in survival over the past few decades, socio-economic gradients are still evident (143,148). Considerable advantage with respect to survival still falls to affluent women (143). This means that many cancer deaths could have been avoided (145). Early diagnosis owing to embedded English screening programmes only partially explains survival inequalities (149). Stage at diagnosis is hypothesised a major contributor to these SES survival patterns, though the role of comorbidities and treatment have also been acknowledged (127). Finally, some recent work suggests that inequality associations

in breast cancer may depend on whether the measurement of SES is at the area or individual level given that reported area level deprivation inequalities were found to be lacking when adjustments for individual effects were considered (150). In England in 1996, there was a 4% deprivation gap in one year breast cancer survival between those women in the least deprived compared to the most deprived locations (142). By 2006, the deprivation gap was still evident, though to a lesser extent (2.6%) (142).

For lung cancers, residing in an area of low SES is associated with shorter survival (104,142,151). Though this socio-economic pattern by low SES has not been replicated in all instances (152,153). As per breast cancer, the level of measure of SES (individual or area) can influence the results. For example, a recent meta-analysis found a weak association in survival for countries using individual incomes measures but no consistent association for education or occupation measures (154). Where SES survival inequalities are seen, diagnosis, treatment and patient factors are often implicated (154). Timeliness of referral and treatment do not always explain SES survival differences, but inequalities in treatment receipt have been observed to do so (104).

1.7 Cancer Treatment

Conventional treatments (surgery, chemotherapy and radiotherapy) have historically dominated cancer treatment guidance (100). Novel anti-cancer therapies (precision medicines, biologicals and immunotherapies) now represent a new era of cancer treatment, owing to our increased understanding of the molecular basis of tumour progression (155). Patients can receive a combination of conventional treatments with/without the inclusion of new novel anti-cancer therapies. Treatment approach depends on the following: cancer type; tumour characteristics (e.g. size, and metastases); the patient (e.g. ability to tolerate cytotoxic therapy or their treatment preferences); and, in the case of some novel anti-cancer therapies, a predictive or resistance biomarker. Whilst these developments have increased the choice of systematic anti-cancer therapies (SACT), no single treatment, which can cure all patients, even in those with similar cancer types, has been discovered (156).

Historically, surgery provided most success at managing early stage, localised tumours with a curative intent (157). It was not until during the second half of the twentieth century that cytotoxic chemotherapies would come to represent a major medical advance in cancer treatment (158). These non-selective drugs were often discovered based on compounds with known cytotoxic activity in similar agents (e.g. the development of alkylating agents from nitrogen

mustard agents used in World War II) (159–161). Chemotherapies effectively disrupt rapid cancer cell division and prevent further tumour growth. This has improved patient outcomes given their large range of applications - though their use is not always curative. Examples include the following: platinum and alkylating agents; anthracyclines; antimetabolites; topoisomerase inhibitors; taxanes; and the vinca alkaloids. Such agents are often limited by their adverse effects and can have long term problems (e.g. alkylating agents are themselves carcinogenic (162) and anthracyclines are linked to cardiac issues (163)). Radiotherapy uses high energy rays or radioactive substances to damage tumour cells, halting their division and growth (164). Radiotherapy is an important conventional therapy, often used alongside other treatment modalities in in both early and late-stage cancers. As its use is wide ranging (curative, neoadjuvant, adjuvant and palliative), Cancer Research UK (CRUK) state that nearly 50% of patients will receive radiotherapy at some point as part of their cancer management (165).

1.7.1 Conventional Treatments for Breast Cancer

Stage (early versus metastatic disease) is important in breast cancer management. For early stage breast cancer (stages I-III), surgery (excision or mastectomy with/without lymph node clearance as appropriate) is first line treatment (166). Further surgery may be needed if radial margins are not clear following initial excision and local recurrence is a thus deemed a risk (166). Breast conserving surgery is prioritised, though breast reconstructions (including a delayed reconstruction) is possible, if desired (166). In patients with predictive (e.g. biomarker status) and prognostic (measurement associated with a clinical outcome e.g. nodal status) risk factors, adjuvant therapy may be required (166). For some patients, adjuvant therapy will compromise endocrine treatments (166). Adjuvant chemotherapy (usually a taxane and anthracycline) can also be used in early stage invasive breast cancer (166). Radiotherapy (whole or partial breast) following breast conserving surgery is also standard practice to improve disease-free survival and reduce recurrence risk (166). Breast surgery is not standard treatment for metastatic disease considered to be incurable and with poor prognosis (167). Where surgery is used in metastatic disease, it tends to be palliative to relieve local bleeding, infection or pain (167). However, with advances in drug treatments, several recent retrospective studies suggest that breast surgery could increase survival in a metastatic instance, though a Cochrane Review concluded that overall the evidence remains uncertain (167). In metastatic disease, chemotherapy is indicated providing consideration of toxicity tolerance. For some patients, adjuvant endocrine therapies will also be an option (168). Finally, radiotherapy use in a metastatic setting tends to be palliative symptom relief (169).

1.7.2 Conventional Treatments for Lung Cancer

In lung cancer, surgery (e.g. lobectomy, broncho-angioplastic surgery, bilobectomy, pneumonectomy) with curative intent can be used on resectable tumours lower than stage IV, if the patient is able to tolerate the surgery (170). Adjuvant treatment with chemotherapy and/or radiation post resection can help reduce the risk of relapse (170). Radiotherapy can also be beneficial, especially in cases of unresectable tumours localised to the chest (170). However, as a significant proportion of diagnoses present at stage IV (this contrasts with breast cancer where established screening helps earlier stage detection), curative intent surgery is limited in use. Despite this, surgery is not contraindicated in stage IV and there is some evidence that it may increase local control on disease progression, though this has not translated into standard clinical practice (171). Rather, the approach in metastatic cancer to achieve disease control and improving quality of life will often be cytotoxic chemotherapy (first line) and palliative radiotherapy (170).

1.7.3 Novel Anti-Cancer Therapies

For this thesis, novel anti-cancer therapies are defined as: precision medicines (molecular targeted therapies); biologicals (anti-angiogenics - drugs blocking blood vessels growth which ordinarily would support tumour growth); and immunotherapies. Other novel interventions such as robotic surgery and stereotactic ablative body radiotherapy (SABR) are not considered.

Novel anti-cancer therapies refer to innovative medicines guided by tumour biology and/or which use the immune system to addresses the lack of specificity in conventional cancer treatments (172). These treatments target one or more of the Hallmarks of Cancer (framework by which the enabling characteristics of malignant cell acquisition is classified) (173) (Figure 1.4. For example, the drug gefitinib (an EGFR inhibitor - EGFRi), targets the mechanisms by which a tumour cell can ordinarily sustain proliferative signalling on account of EGFR dimerisation and the consequent downstream cell signalling processes which ensue and moderate cell proliferation and migration (174). Novel anti-cancer therapies can provide tailored, precise and accurate health interventions which maximise patient benefit (175).

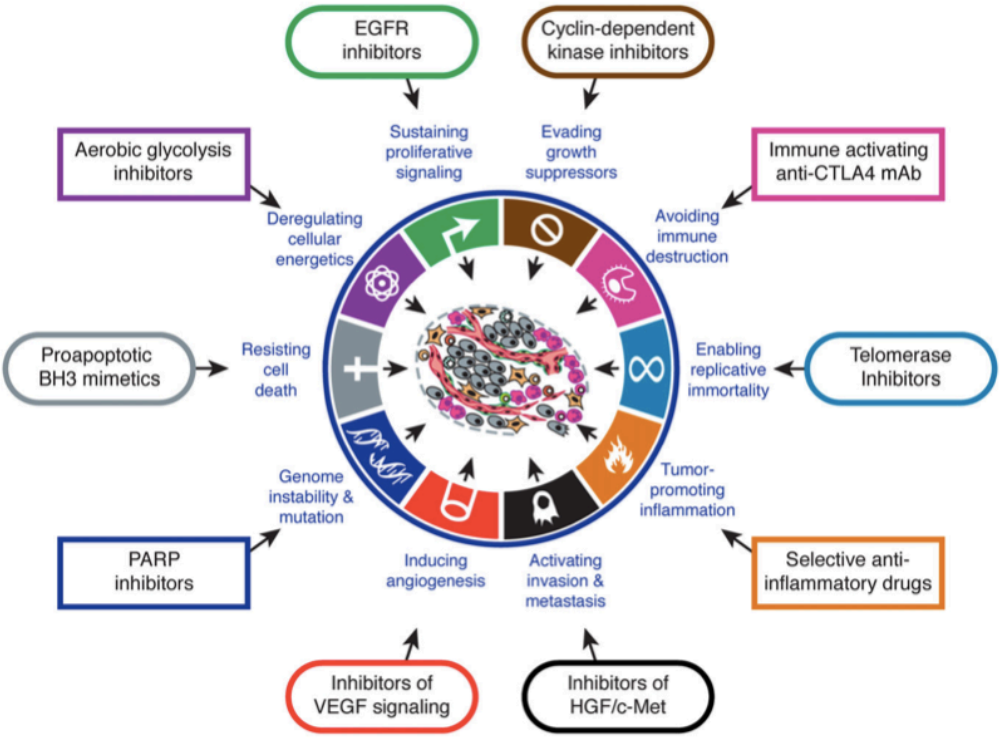


Figure 1.4 The hallmarks of cancer (Taken from Hanahan and Weinberg, 2011) (173).

1.7.4 Precision Medicine

Precision medicine is an emerging term that refers to customisable treatments which consider variability in genes, environment and lifestyle factors shared by a sub-group of patients (172). Tailored pharmacotherapy, whilst important, is not the sole component of precision medicine (176). Rather, the term also encapsulates: (i) disease prevention through risk prediction; (ii) differential diagnosis and timely identification; and (iii) optional treatment (176).

The term precision medicine first appeared in the late 1990s, though many consider 2011 as the date when the precision medicine grew in popularity and become synonymous with "stratified", "individualised" and "personalised medicine" (now superseded by a preference for "precision") (177,178). Precision medicine reflects the evolution of care, away from historically reactive and disease-based approaches towards a more pro-active "P4" medicine vision (predictive, preventive, personalised and participatory medicine) (177). Targeted treatment is not a new concept (it has long been realised that patients are unique). However, new diagnostic technologies and biologically targeted therapies make the promise of a "targeted" approach seem increasing feasible (179,180).

Oncology has particularly benefited from the introduction of precision medicine as observed by the rapid rise, over the past two decades, in the use of cancer molecular diagnostics

(100,181). The basic idea of precision cancer medicine is that patients are assigned therapies which their tumour is mutated for and hence likely to respond too (182,183). Tumour mutation drivers of interest (identified through predictive biomarker testing) include the following: oncogene addiction; loss of tumour suppressor genes; synthetic lethality associated with DNA damage repair; and antibody drug conjugate targets.

For this thesis, precision medicine references molecular targeted therapies, targeting either oncogene addiction or synthetic lethality with activity restricted to tumours with appropriate biomarker status. Hormone therapies are excluded from the definition, even though anti-oestrogens (endocrine therapies for breast cancer) are considered, by some, to be the first cancer targeted therapy (ER positivity is an indicator for tamoxifen or aromatase inhibitors) (159,184). This is because such treatments are well established; they have been in place since the 1970s and their patterns of utilisation are well studied (185).

1.7.5 Biomarker Tests

The National Cancer Institute (NCI) defines a biomarker as "a biological molecule found in blood, other bodily fluids, or tissues that is a sign of a normal or abnormal process, or of a condition of disease" (186). Biomarkers in cancer differentiate patients with tumour mutations of interest and thus provide a helpful decision-making tool in determining those patients likely to respond and benefit from targeted therapy. In the context of cancer treatment, there are multiple biomarkers (e.g. predictive, prognostic, pharmacodynamic, markers of tumour response). Prognostic and predictive biomarker differences are discussed further below.

Prognostic Biomarkers: These evaluate overall tumour outcomes independent of therapies given e.g. recurrence probability (155). Prognostic markers can aid selection of patients for treatment (considering relapse or death risk along with cases where toxic treatments should be omitted) but such tests do not directly predict treatment response (187). MammaPrint (70 gene prognostic assay) and Oncotype DX (21 gene multiplex test assay) are examples of prognostic biomarkers, evaluating both breast cancer recurrence, risk and prognosis (155).

Predictive Biomarkers: These define the likelihood of a pharmacological response (whether the tumour is treatment sensitive or drug resistant), which is used to guide patient stratification and sub-grouping (188,189). Predictive biomarker presence is often a pre-requisite for access to many, but not all, precision medicines (155). Examples include: EGFR sensitising mutation to guide EGFRi use in NSCLC or HER2 positivity to guide trastuzumab use in breast cancer (189).

Immunotherapy use can also benefit from predictive biomarkers (e.g. Programmed Cell Death Protein Ligand 1 (PD-L1) tumour proportion score), though expression does vary with tumour site and a significant number of patients with PD-L1 negative tumours can still benefit from PD-L1 inhibitors (190).

Biomarker test results can be binary (presence or absence) or a percentage score (extent of expression). Some biomarkers are both prognostic and predictive (155). For example, Breast CAncer Gene 1 (BRCA1) is a prognostic marker when determining prognosis but predictive when guiding chemotherapy (breast cancer) or Poly (ADP-Ribose) Polymerase (PARP) inhibitor initiation (ovarian cancer) (155). Pharmacological decision making is easier in monogenic conditions where single biomarkers infer treatment choice (191). However, many cancers are multifactorial and have many biomarkers, hence testing does not necessarily simplify treatment decision making (191). Additionally, not all cancers have molecular drivers identified (192). Examples of common predictive biomarker tests for different cancers are outlined in Table 1.2.

Table 1.2 Common predictive biomarker tests for a selection of solid tumour cancers

Cancer	*Predictive Biomarker Test*
Breast	ER
	HER2
	PR
Colorectal	BRAF
	KRAS
	MLH1
	MSH2
	MSH6
	NRAS
	NTRK1
	NTRK2
	NTRK3
	PMS2
	POLD1
	POLE
Lung	ALK
	ELM4-ALK
	BRAF
	EGFR
	MET
	NTRK1
	NTRK2
	NTRK3
	PD-L1
	KRAS p.(G12C)
	RET
	ROS1
Melanoma	BRAF
	KIT
	NRAS
	NTRK1
	NTRK2
	NTRK3
Ovarian	BRCA1
	BRCA2
	SMARCA4
	NTRK1
	NTRK2
	NTRK3

Abbreviations: ALK: Anaplastic lymphoma kinase; BRAF: V-raf murine sarcoma viral oncogene homolog B1; BRCA1/2: BReast CAncer gene 1/2; EGFR: Epidermal growth factor receptor; ER: Oestrogen receptor; HER2: Human Epidermal Growth Factor Receptor 2; EML4-ALK: Echinoderm microtubule-associated protein-like 4 anaplastic lymphoma kinase; KIT: Feline sarcoma viral oncogene v-kit; KRAS: Kirsten rat sarcoma virus; KRAS p.(G12C): Kirsten rat sarcoma virus glycine-to-cysteine substitution at codon 12: Kirsten rat sarcoma virus; MET: Mesenchymal-epithelial transition factor; MLH1: MutL homolog 1; MSH2: MutS homolog 2; MSH6: MutS homolog 6; NRAS: Neuroblastoma ras viral oncogene homolog; NTRK1/2/3: Neurotrophic tropomyosin-receptor kinase 1/2/3; PD-L1: Programmed cell death-ligand 1; PMS2: PMS1 homolog2, mismatch repair system component; POLD1: DNA polymerase delta 1, catalytic subunit; POLE: Polymerase epsilon catalytic subunit; PR: Progesterone receptor: RET: Rearranged during transfection; ROS-1: Receptor tyrosine kinase 1; SMARC4: SWI/SNF related, matrix associated, actin dependent regulator of chromatin, subfamily A, member 4.

1.7.6 Biomarker Tests for Breast Cancer

In England, testing for endocrine receptors and HER2 at the time of breast cancer histopathological diagnosis is routine (Table 1.2); immunohistochemistry (IHC) and/or fluorescence in situ hybridisation (FISH) techniques are used (77,97). For the HER2 positive (HER2+) breast cancer subtype, HER2 is an important biomarker determining suitability for trastuzumab or other anti-HER2 medicines. Testing for the ER (and less so for the PR) can also be useful, not only for subtype classification and prognosis calculation (endocrine receptor positive tumours tend to have better outcomes) but also to determine the need for endocrine therapy (155). Screening for new, additional breast cancer biomarkers (e.g. phosphatidylinositol 3-kinase (PI3K) prior to alpelisib use) is now also approved by the Medicine and Healthcare products Regulatory Agency (MHRA).

1.7.7 Biomarker Tests for Lung Cancer

Mutations in specific biomarkers are rare individually (bar kristen rat sarcoma virus (KRAS)), but as up to 60% of adenocarcinomas may contain a driver mutation, testing and directing therapy is worthwhile (98). Current UK and international guidelines recommend non-squamous NSCLC patients undergo testing for numerous biomarkers (Table 1.2) where appropriate to identify suitable first line therapy (193–195). There is also a 10 day "molecular testing standard" recommendation by NHS England's National Optimal Lung Cancer Pathway - though success meeting this target is poor (193). Of particular interest to this thesis are EGFR and anaplastic lymphoma kinase (ALK) mutations - and their associated inhibitor drugs. This is because ALK and EGFR targeted treatments were licenced, and their use was well established during the time frame of data analysis in Chapters 3 and 4. In EGFR, exon 19 (DELL19) and exon 21 Leu85Arg substitutions are most common (100). In the UK, recent work has identified a 10.2% EGFR mutation rate (196). ALK mutations are thought to occur in 5% of metastatic NSCLC and provides a further established NSCLC druggable mutation (197). NICE guidance also mentions receptor tyrosine kinase 1 (ROS1) and neurotrophic tropomyosin-receptor kinase (NTRK) fusions thought to occur in 1% and less than 1% of metastatic NSCLC respectively (197). There is some integration of biomarker testing in an immunotherapy use setting. For example, patients with a PD-L1 tumour proportion score of 50% or greater may benefit from an ICI (usually pembrolizumab) (197) which blocks PD-L1 to restore T-cell mediated immunity to stop tumourigenesis (72). It is anticipated that further molecular targets will be integrated into clinical practice in the future e.g. mesenchymal epithelial transition factor receptor (MET) exon 14 skipping mutations (197).

1.7.8 Molecular Targeted Therapy

There is no universal consensus on the definition of a targeted therapy and there is debate over the concept given that all treatments have a target (known or otherwise) (159,160). The NCI therefore defines such therapies as "drugs or other substances to target molecules involved in the growth and spread of cancer cells" (198). The basic principle of a targeted therapy is thus targetable, precision differentiation of healthy cells from cancerous ones (100). Molecular targeted therapies block the growth and spread of cancer by inhibiting the activity of mutated or over-expressed oncogene(s) (oncogene addiction) or a pathway that the tumour has become overly reliant on (e.g. a compensatory mechanism for other molecular abnormalities - synthetic lethality) (158,199). There are many classes of targeted drugs. Molecular targeted therapies of relevance to HER2+ breast cancer are shown in Table 1.3 (page 30) and NSCLC in Table 1.4 (page 34).

1.7.9 Biological/Biologic Therapies

There is no established definition of a biological therapy in relation to cancer that accurately describe treatment modalities of interest which are not captured by molecular targeted therapy, immunotherapy, or cytotoxic chemotherapy definitions. To help differentiate such therapies in this thesis, a biological/biologic therapy was thus defined as a specific targeted therapy which has no associated predictive biomarker status included in the licence. For this thesis, biologicals relate to anti-angiogenics and/or any broad literature reference (defined or otherwise) to a "targeted biologic/biological". An example of this type of treatment, used in both breast and lung cancers is bevacizumab, a MAB targeting vascular endothelial growth factor (VEGF) and nintedanib (kinase inhibitor) in NSCLC (Tables 1.3 and 1.4).

1.7.10 Immunotherapy

An immunotherapy is defined as a treatment that helps "the immune system recognise and attack cancer cells" (200). Immunotherapies stimulate the immune system's natural, adaptive capacity to develop durable memory and use this to assist cancerous cell recognition and destruction, usually through mechanisms that target immune resistance or by modulating T-cell functions (190,201). Immunotherapies do represent a form of targeted treatment, either as a single agent or when offered in combination therapies, but as they target the immune system/evasion of immune surveillance as opposed to directly impacting cancer signalling, they will be referred to throughout with their own classification. These treatments were of interest to this thesis because they constitute a new targeted treatment line and one which has received

considerable interest (as observed by the exponential growth in their development in the past decade) (53). It is therefore important that their associations with SES are understood.

Immunotherapies of interest are PD-L1 immune checkpoint inhibitors (ICIs) which can be used in NSCLC. ICIs have shown promise at effector T-cell inhibition which helps preserve the tumour killing function of cancerous cells (190) or through preventing the activation of choline transporter-like protein 1-4 (CTL1-4) and this helps restore immunity (190). Immunotherapies can have associated biomarker testing (e.g. PD-L1 tumour proportion score) and this can be useful for therapy guidance, though there is some instance of ICI benefit even when PD-L1 scores are low (190). NSCLC has benefited from immunotherapy additions into treatment guidelines (202,203). There is now licenced immunotherapy use in breast cancer; but at present this is limited to TNBC only (204). However, in time, it is anticipated that immunotherapy use will widen and include breast cancer sub-types other than just TNBC.

1.7.11 Precision Medicine in HER2+ Breast Cancer

Precision medicine use in HER2+ breast cancer involves anti-HER2 molecular targeted therapies and/or biological therapy (Table 1.3). These novel anti-cancer therapies often form one strand of a treatment approach which may also include: surgery, chemotherapy, radiotherapy and/or, endocrine therapy. Treatment options and sequencing vary on an individual basis (patient choice and tumour characteristics) (205). Staging is important in breast cancer treatment, with early-stage patients usually receiving surgery first (Figure 1.5). This is not the case in metastatic breast cancers. Most women will receive one or more chemotherapy drugs in addition to trastuzumab (205). Some HER2+ breast cancers may also receive an additional anti-HER2 therapy. Combination therapies are often offered with the aim of achieving a higher synergistic response e.g. trastuzumab and pertuzumab (100). Pertuzuamb, a humanised MAB, is effective when used in combination with trastuzumab as it binds to domain II of HER2 (opposed to domain IV) and this inhibiting ligand induced HER2/HER3 dimerisation, which trastuzumab use alone only has a minor effect on (206). The MAB trastuzumab should be offered to patients with a HER2+ status after consideration of: comorbidities; prognosis; drug toxicity and patient tolerability; body mass index (BMI); menopause status; and tumour grade (97,184). Trastuzumab can be offered in combination with surgery and radiotherapy, where appropriate (97) and is cautioned in patients with a cardiac history, although cardiac monitoring every three months can mitigate this risk (97). For a standard patient with HER2+ breast cancer it would be reasonable to anticipate a treatment approach of: systemic anti-cancer therapy/ies administration 3-6 months post diagnosis

(including anti-HER2 therapy to start before, during or after chemotherapy and lasting for around 12 months), with surgery occurring before or after chemotherapy.

Table 1.3 Novel anti-cancer therapies licenced for HER2+ breast cancer in England

Novel Anti-Cancer Therapy	***Drug Class***	***Drug***	***Associated Predictive Biomarker Test***
Biological Therapy	Anti-Angiogenics	Bevacizumab	None
Molecular Targeted Therapy	Anti-body Drug Conjugates	Trastuzumab Emtansine Trastuzumab Deruxtecan*	HER2
	MABs	Trastuzumab Pertuzumab	HER2
	Small Molecule TKIs	Lapatinib Neratinib Tucatinib	HER2

Abbreviations: HER2: Human Epidermal Growth Factor Receptor 2; MABs: Monoclonal antibodies; TKI: Tyrosine kinase inhibitors.
*Deruxtecan is a conventional chemotherapy

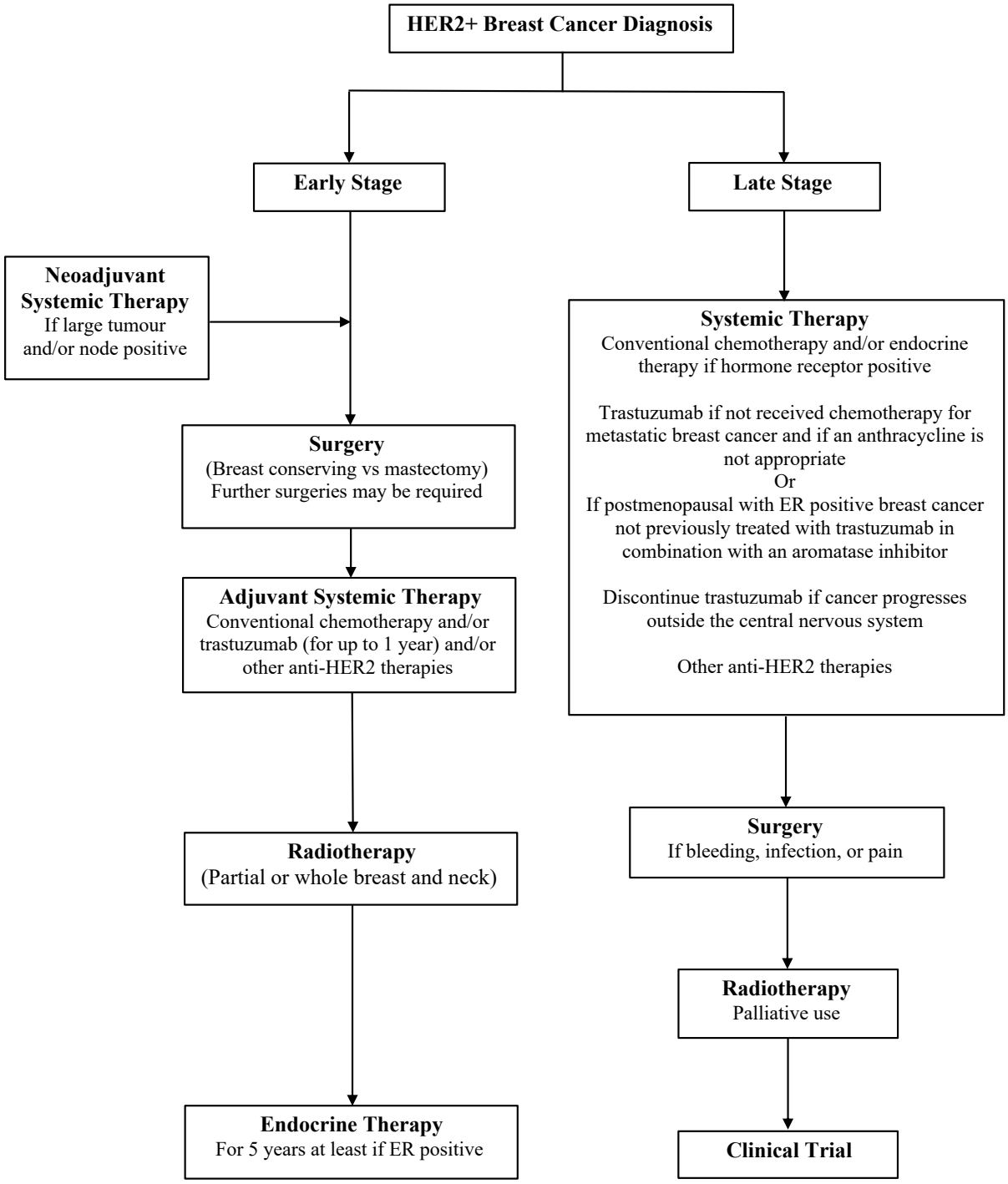


Figure 1.5 A simplified flow diagram to illustrate treatment ordering decisions for a "standard" patient with HER2+ breast cancer by stage at diagnosis. Treatment is anticipated to last approximately one year. Treatment ordering can vary on a case-by-case basis. For example, surgery can be started before, during or after chemotherapy. Additionally, each treatment step will not be considered appropriate in all instances.

1.7.12 Precision Medicine in NSCLC

Patients with stage I-III NSCLC are generally treated with curative intent using surgery, radiotherapy and/or (often in a combination treatment approach). Personalised therapies are increasingly being utilised in NSCLC and are now licenced in the adjuvant setting which extend their initial utilisation in advanced disease only (207). Immunotherapy can be used in both early and metastatic NSCLC. In early stage and locally advanced NSCLC, neoadjuvant ICIs plus chemotherapy can increase tumour response rate, whilst adjuvant ICI use leads to increased disease-free survival in PD-L1 positive tumours (208). In metastatic NSCLC, ICIs can also be used, especially in the absence of other oncogenic tumour drivers e.g. pembrolizumab is a standard first line treatment option in EGFR- and ALK- tumours with a PD-L1 score greater than 50% and with no contraindications for immunotherapy use (209,210). Tumours with stage IIIB or IV disease and with a driver mutation present will also be candidates for novel anti-cancer therapy. At present in England treatment guidance is available for EGFR, ALK and ROS1 non-squamous NSCLC cancers. NTRK fusion-positive inhibitors are licensed too but the point at which they are used in treatment guidance is less clear. Additionally, in early-stage NSCLC, adjuvant osimertinib is also approved for use in patients with stage IB to IIIA NSCLC, whose tumours have EGFR exon 19 deletions or exon 21 L858R substitution mutations and have undergone complete tumour resection (208).

As the sequence of conventional treatments is less important than in breast cancer, and as precision medicines of interest are generally contained to later stage disease, only the NICE flow diagram for novel anti-cancer treatment decision making following tumour mutation driver identification (211) is shown here (Figure 1.6). For patients without biomarker susceptible tumours, chemotherapy remains the treatment mainstay (212). As the NSCLC population is heterogenous, there is no “standard” treatment duration for a “typical” NSCLC patient. Time on treatment depends on several patient and clinical factors. For example, for funded ICI in stage IV patients, treatment duration can be up to two years or until loss of clinical benefit. Targeted therapies are typically used until progression and beyond and these timescales may vary considerably (from a few months to a few years).

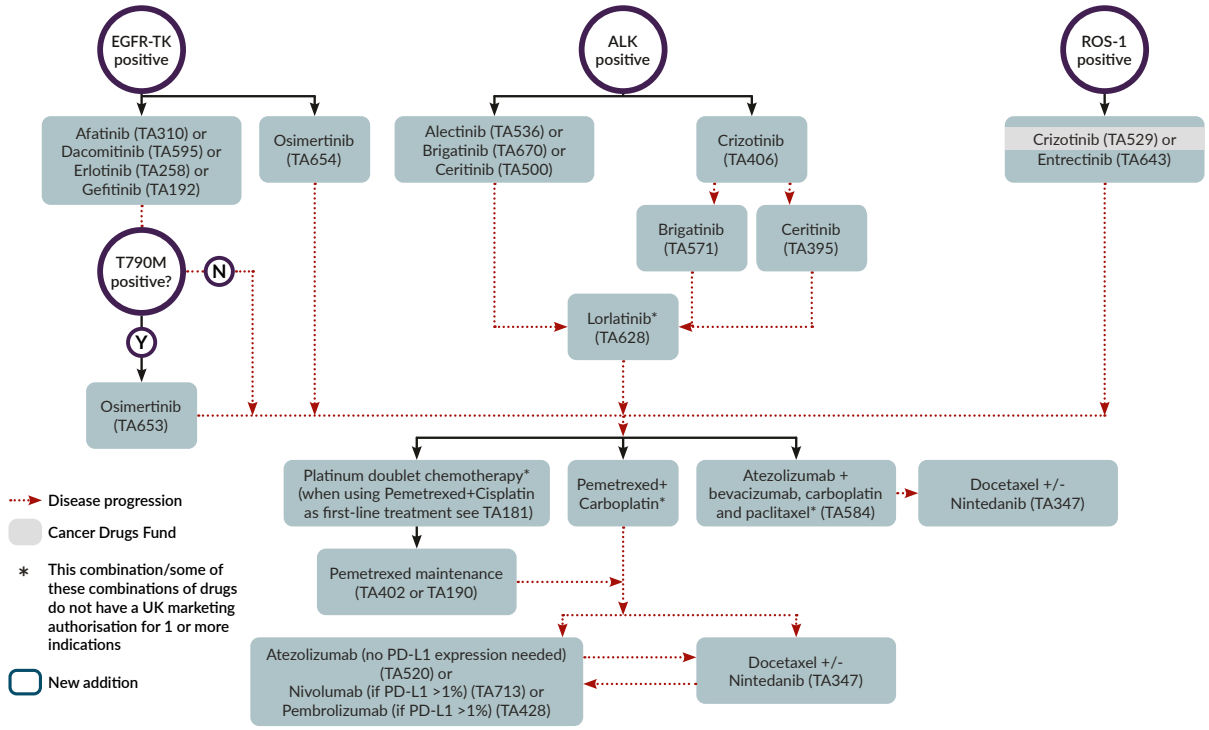


Figure 1.6 NICE guidance for non-squamous NSCLC novel anti-cancer therapy treatment sequencing in England following identification of a tumour mutation driver (211). T790M mutation confers TKI resistance. Abbreviations: ALK: Anaplastic lymphoma kinase; EGFR-TK: Epidermal growth factor receptor tyrosine kinase; PD-L1: Programmed death ligand 1; ROS-1: Receptor tyrosine kinase 1; TA Numbers e.g. TA310 refer to NICE technology appraisal guidance number; TKI: Tyrosine kinase inhibitor; T790M: Gatekeeper mutation of the epidermal growth factor receptor;

Table 1.4 Novel anti-cancer therapies licenced for NSCLC in England

Novel Anti-Cancer Therapy	***Drug Class***	***Drug***	***Associated Predictive Biomarker Test***
Biological Therapy	Anti-angiogenics	Bevacizumab Nintedanib	None
Immunotherapy	PD-L1 ICIs	Atezolizumab Durvalumab Nivolumab Pembrolizumab	PD-L1/None
	Anti-CTLA4	Ipilimumab	None
Molecular Targeted Therapy	ALKi	Alectinib Brigatinib Ceritinib Crizotinib Lorlatinib	ALK/ELM4-ALK
	BRAF Inhibitors	Dabrafenib Trametinib	BRAF V600
	EGFRi	Afatinib Dacomitinib Erlotinib Gefitnib Osimertinib	EGFR
	MET Inhibitors	None Licenced	MET
	NTRK Inhibitors	Entrectinib Larotrectinib	NTRK1/NTRK2/NTRK3
	RET Inhibitors	Selpercatinib	RET
	ROS1 Inhibitors	Ceritinib Crizotinib Entrectinib	ROS1

Abbreviations: ALK: Anaplastic lymphoma kinase: ALKi: Anaplastic lymphoma kinase inhibitors: Anti-CTLA4: Anti cytotoxic T-lymphocyte-associated protein 4; BRAF V600: V-raf murine sarcoma viral oncogene homolog B1 V600; EGFR: Epidermal growth factor receptor; EGFRi: Epidermal growth factor receptor inhibitors: EML4-ALK: Echinoderm microtubule-associated protein-like 4 anaplastic lymphoma kinase; ICIs: Immune checkpoint inhibitors; KRAS G12c: Kirsten rat sarcoma virus glycine-to-cysteine substitution at codon 12; Kirsten rat sarcoma virus; MET: Mesenchymal-epithelial transition factor; NTRK1/2/3: Neurotrophic tropomyosin-receptor kinase 1/2/3; PD-L1: Programmed cell death-ligand 1; RET: Rearranged during transfection; ROS-1: Receptor tyrosine kinase 1.

1.7.13 Socio-economic Associations in Cancer Treatment

Utilisation of cancer treatment has been shown to vary by SES, even though there is no real reason why this should occur - especially in publicly funded healthcare systems where care is (or should be) provided based on need (122). Several studies have hypothesised that SES differences in certain cancer treatments may also explain survival and mortality variations (126,213). Prioritising treatment access has thus been a focus in many cancer care improvement initiatives (214,215).

There is evidence for SES inequalities in breast cancer treatment. For example, studies have shown reduced utilisation in: breast conserving surgery (92,216–219); neoadjuvant chemotherapy (220); and radiation therapy (92) for patients with a lower SES. Similar reduced likelihood of treatment for patients with a low SES have also been observed in lung cancer both for surgery (219,221–223) and chemotherapy (221,223,224), though not always for radiotherapy (221).

1.8 Thesis Rationale

It is reasonable to hypothesise that new medical treatments introduced to improve overall population health would do so. This has been a major promise of precision medicine - improved outcomes for all, curtailing engrained health inequalities and advances in precision public health (172,225,226). Such promises are plausible given that novel anti-cancer therapy prescribing reduces the uncertainty associated with conventional clinical decision making (227). This is because genomic medicine ties clinical decisions to the results of a test for a specific gene aberration, hence the impact for bias and stereotyping, especially on the basis of socio-economic characteristics should be minimised (227).

On the other hand, some scholars take a more cautionary approach, arguing that the introduction of new medical advances may potentially widen inequalities further (58,228). Such a phenomena is termed an “intervention generated inequality” (IGI) (56). Such proponents argue that genes, rather than acting in isolation, are only one part of complex set of interactions influencing health (i.e. epigenetic processes, behaviour and the environment) (57,183,191,229). Proponents do not necessarily deny the benefit of precision oncology in defined cohorts to improve health outcomes, however they do raise concerns that proposals for small scale and selective applications of treatment will narrow pre-existing and engrained cancer inequalities in conventional cancer care (52,53). Others propose more of a middle ground. That is, new interventions result in a temporary widening of inequalities for low SES patients after their

introduction due to preferential intake by the most advantaged (the Inverse Equity Hypothesis) (230). However, over time, this inequality narrows as the intervention "trickles down" into less advantaged groups and becomes standard practice.

At present therefore, it remains unknown whether novel anti-cancer therapies are subject to similar inequality drivers as conventional treatment. The scholarly debate on health inequalities in precision medicine utilisation would benefit from an analysis of the real-world utilisation of these novel, emergent anti-cancer therapies as recorded in big data (large, complex datasets such as those routinely collected as part of electronic health care records (EHR)). This work is significant because should inequalities be observed, this may have important consequences on cancer outcomes such as mortality and survival. This work is also timely given the speed of change within cancer treatments guidelines of late and the implications this has on ensuring that fair treatment for all adapts to changing treatment circumstance.

1.9 Aim and Scope

The aim of this thesis is to establish whether (and if so - to what extent) there are socio-economic inequalities in the utilisation of novel anti-cancer therapies and their associated predictive biomarkers.

Limits to scope have been discussed throughout this chapter and are summarised again. This thesis will focus only on solid tumour cancers, predictive biomarker tests, and a definition of a targeted therapy which does not reference endocrine treatments.

The specific research objectives are as follows:

- Systematically review the existing evidence exploring associations between SES and predictive biomarker testing utilisation.
- Systematically review the existing evidence exploring associations between SES and novel anti-cancer therapy utilisation.
- Determine the association of SES (measured in terms of the IMD income domain of area of residence at diagnosis) on trastuzumab and conventional treatment (breast cancer directed surgery and chemotherapy) utilisation for a HER2+ breast cancer population in England.
- Determine the association of SES (measured in terms of the IMD income domain of area of residence at diagnosis) on novel anti-cancer therapy utilisation for a NSCLC population in England.

1.10 Thesis Overview

The remainder of this thesis is structured in four Chapters. Chapters synopses are described below to provide further details on their content.

Chapter 2: Describes a systematic review and meta-analysis of observational studies reporting data on SES and novel anti-cancer therapies as well as their associated predictive biomarker test utilisation. The chapter identifies, appraises, analyses, and synthesises the results of previous peer reviewed publications meeting the review's inclusion criteria.

Chapter 3: Details a population-based observational study of secondary data obtained from the English National Cancer Registry Database (NCRD) and systemic anti-cancer therapies (SACT) dataset. The Chapter determines associations of trastuzumab utilisation on a cohort of HER2+ women with their deprivation level. Methods for obtaining the data are detailed along with an analysis of the results.

Chapter 4: Provides a further population-based observational study of English NCRD and linked SACT secondary data. The Chapter determines associations of any novel anti-cancer therapy utilisation on a cohort of NSCLC patients with their deprivation level. Methods are detailed along with an analysis of the results.

Chapter 5: Summarises the thesis by integrating the findings of preceding chapters. The Chapter records the main research findings in the context of the original aims and objectives. Strengths and limitations are noted along with implications for future research, policy, and practice. Closing remarks conclude the work.

1.11 Conclusions

This chapter provided a background on the move towards stratified pharmacotherapy in cancer care. Targeted, novel anti-cancer therapies have grown in popularity and offer opportunity for improved outcomes (reduced side effects profiles and increased progression free survival). However, it is less clear if these treatments, as with conventional cancer therapies, are subject to socio-economic inequalities in utilisation. To begin answering this question, Chapter 2 now examines whether novel anti-cancer therapy utilisation, already outlined in the published literature, varies by SES.

Chapter 2. Are there Socio-economic Inequalities in Utilisation of Predictive Biomarker Tests and Novel Anti-cancer Therapies? A Systematic Review and Meta-analysis

2.1 Introduction

Chapter 1 summarised the evolving nature of cancer treatment considering recent medical advances stemming from a growing appreciation of the role of tumour biology in carcinogenesis. It also highlighted that cancer treatment utilisation inequalities have been reported for conventional treatments. It ended with a discussion on whether the move towards personalised, precision treatments, which identify sub-populations to target for treatment success, will improve or worsen current socio-economic inequalities in treatment utilisation. To begin addressing this question, an analysis of the state of current, real-world novel anti-cancer therapy utilisation by SES is needed. At the time of conducting this work, no such analysis existed; this Chapter therefore serves to address this knowledge gap.

Chapter Aim:

Systematically review the existing evidence exploring associations between SES and associated predictive biomarker test and/or novel anti-cancer therapy utilisation.

Objectives:

1) Conduct a systematic review to summarise the existing state of knowledge regarding possible associations between: (i) predictive biomarker tests; and/or (ii) novel anti-cancer therapies and a measure of SES.

2) Perform a meta-analysis to combine data reporting the likelihood of :(i) predictive biomarker tests; and/or (ii) novel anti-cancer utilisation by SES.

This Chapter details the methods and results of this analysis. The work undertaken for this chapter has been published (see Norris *et al*., 2020; Appendix 2.1) (231). Hence the information presented here reflects the original published text, with amendments, where necessary, to provide supplemental information for the flow of this thesis.

2.2 Background

Individuals with a lower SES are less likely to receive conventional cancer treatments and this may contribute to poorer cancer outcomes in this group (124). Increasingly systemic treatments targeted at cancer biology (e.g. TKIs and MABs) are being integrated into routine cancer clinical care. These agents are expensive (immunotherapy can cost in US dollars, $100,000 per patient annually) and may only have efficacy in selected sub-populations (232). Hence stratifying patients by molecular pathology to predict the likelihood of tumour response and adjusting therapy accordingly is now routinely recommended (see, for example, (233)). This move towards novel anti-cancer therapies is reflected in the cancer drug development pipeline; for example, in 2019, 450 new cancer drug candidates (representing some of the most promising drugs on the market) were immunotherapies (234) and, by 2021, these treatments provided the largest area of cancer research (235).

Socio-economic inequalities in novel anti-cancer therapy utilisation remains largely unexplored. A meta-analysis by Martin *et al.* (2018) has previously reported pooled OR data on trastuzumab uptake by SES which concluded equivocal results (236). However, this study only focused on one novel anti-cancer therapy (trastuzumab) and synthesised only a few studies (n = 5 studies; 4,294 patients) (236). Hence further work is still needed as there is speculation that using molecular information to target cancer treatment may potentially provide a solution to current treatment inequalities (237). However, others argue that novel cancer therapies, because of their cost, disproportionately favour those with more resources and, therefore, may widen inequalities further (57,238,239).

As novel anti-cancer therapies and their associated predictive biomarker tests offer opportunities for increased tumour response, reduced adverse effects and improved survival, it is important to understand whether there are inequalities in their utilisation (240,241). This systematic review and meta-analysis identified and integrated the existing research to investigate the relationship between SES and utilisation of novel anti-cancer therapies and their associated predictive biomarker tests.

2.3 Materials & Methods

The review was registered with the international database of prospectively registered systematic reviews, PROSPERO (CRD42019140016) and is reported according to the version of the

Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) in use at the time that the review was conducted (242) (Appendix 2.2)[1].

2.3.1 Search Strategy and Study Selection

Searches were performed in seven databases (MEDLINE, Embase, Scopus, CINAHL, Web of Science, PubMed and PsycINFO) for articles published between January 1998 and December 2019. These databases were chosen for their general scope; in combination there were considered likely to contain the types of studies the research question was trying to identify. This time period reflected the licensing and approval of trastuzumab in the USA - considered by Huang *et al.* (2014) as a crucial time marker in the precision therapy field (158). Hence it was considered unlikely that searching prior to 1998 would yield any further additional articles of interest. Grey literature (documents not controlled by commercial publishing) was not included in the search due to the absence of peer review in such texts along with difficulty systematically searching for and retrieving this information. Novel anti-cancer therapies of interest included: targeted therapy (targeting either oncogene addiction or synthetic lethality with activity restricted to tumours with appropriate biomarker status); biologics (where no predictive biomarker is included in the license); and ICIs. Therapies targeting endocrine receptors (i.e. tamoxifen and aromatase inhibitors) were excluded as these agents have been in use since the early 1970s and their use is well documented (185).

Search terms covering SES, test, and therapies were developed. Reference lists of eligible articles were also reviewed for relevant studies. A full search strategy is available in Appendix 2.3. The search strategy for the review was kept deliberately broad. This was because prior to undertaking the review, the types of cancers, novel anti-cancer therapies and measures of SES already reported on were unknown. Scoping searches thus formed an important piloting stage to narrow the review's focus and to ensure the retrieval of key papers within the results. At the time of the search, there was no universal definition/terminology of a novel anti-cancer therapy nor a licensed drug list to draw search terms from (244). Treatment search terms were initially generated from the European Medicines Agency's (EMA) list of approved solid tumour cancer drugs (downloaded on 02/01/2019), and amended following cross referencing with any novel anti-cancer drug classification in the then current British National Formulary (BNF) (Edition

[1]PRISMA statement (2009) (Appendix 2.2) (242) as opposed to the updated PRISMA (2020) statement (243) was used was as this was the version in use at the time that the review was both undertaken and published.

76) (245,246). Clinical judgment, exercised by the team (AG, RN & AT) (Appendix 2.4)[2] refined the drug search terms of interest. All EMA clinical trials drugs were omitted as it was concluded unlikely that such therapies would be recorded in the national English cancer registry and there was a desire to have some continuity of scope between this review and later analyses presented in this thesis.

The inclusion criteria for full text papers (which would be expected to have been subject to more scrutiny as part of the peer review process than to conference proceedings and abstracts), published 1998 onwards and written in English (access to translation services were not available) were determined in terms of the PICOS (Population, Intervention, Comparison, Outcome and Setting) framework, as follows:

Population: Any solid tumour cancer diagnosis at any age or sex. Studies meeting all other criteria but reporting no denominator population were included but synthesised separately.

Intervention: Utilisation of either a predictive biomarker test or novel anti-cancer therapy (for definitions see Chapter 1, Sections 1.7.3 and 1.7.5). Any studies reporting novel anti-cancer therapies administered with an adjuvant (e.g. chemotherapy) were eligible as long as it was clear how many patients utilised the novel anti-cancer therapy. Only predictive biomarker tests of pharmacological response to targeted treatment were included. Dual prognostic and predictive biomarkers were only included when it was possible to establish that the data reported the biomarker in that instance was for its predictive not prognostic capacity (see Section 1.7.5 for definition differentiation).

Comparison: It was not a requirement that a comparator was reported but where noted, the following comparator details were extracted: a clinical alternative, no novel anti-cancer therapy and/or predictive biomarker tests or no treatment.

Outcome: This referred to the utilisation data for extraction as reported by a SES measure (e.g. percent of persons living below the poverty line, median household income). Studies reporting utilisation by only an average measure of SES were included but were synthesised separately.

[2]Refers to the initials of the authors as per the published review paper (Appendix 2.4).

Setting: This could be either a retrospective or prospective observational design (including randomised controlled trials analysed as observational cohorts). Full inclusion criteria are listed in Appendix 2.5.

Screening of titles and abstracts was conducted by one author (RN) only. All articles selected for full text review were independently checked by a second author (AT). Disagreements were discussed and if necessary, resolved with a third author (LS). Agreement between reviewers (RN & AT) was excellent ($\kappa = 0.93$) (247).

Study selection is reported throughout this review by (i) the number of papers and (ii) the number of studies meeting the PICOS. This is to account for the fact that multiple publications reporting identical or heavily overlapping study populations were found.

2.3.2 Data Extraction and Quality Assessment

Data was extracted by one author (RN) and checked by another (RD). Disagreements were resolved through discussion with the review team (AG, AT, LS & RD). In instances of missing or inconsistent data, study authors were contacted by email. Where there was no response to an initial or follow up email within six weeks, data was documented as not reported, or the paper excluded. In the event of multiple publications reporting identical or heavily overlapping study populations (e.g. same registry, cancer, stage, age group, and time period) only one paper was selected for reporting in the characteristics of included studies data table, though all other papers were still listed (bolded reference refers to the paper from which the data are extracted). A decision-making hierarchy was used to determine from which paper to extract this data. This was as follows: (i) extract data from the earliest dated (by year) overlapping study publication; (ii) where there was more than one publication reporting identical information from the same year, data extraction then prioritised the publication reporting an income-based socio-economic measure; and (iii) where there was more than one paper from the same year reporting an income based socio-economic measure, data extraction prioritised the paper reporting multiple socio-economic measures. The rationale for this was that most studies reported an income measure and monetary measurements can be at a more specific individual/household level as opposed to a broader area level one. Additionally, as targeted treatments are costly, a patient's ability to pay would be important in a non-publicly funded healthcare system. If more than one multivariable analysis was reported in the paper, information was extracted from the most comprehensive adjusted model.

Data was extracted on: author(s); publication year; country; data source; number in study population; cancer diagnosis time frame; patient age(s), cancer stage and registry coverage; all socio-economic measure(s) and units listed; numbers receiving predictive biomarker test/novel anti-cancer therapy, overall and by socio-economic group (numerator and, where available, denominator); comparator(s) (where appropriate); and measures of association for not receiving testing/treatment by SES (e.g. odds ratios (ORs), 95%, confidence intervals (CIs) and p values).

All eligible studies (including those without a denominator population or which only reported an average measure of SES) were quality appraised using a quality appraisal tool derived from The International Society for Pharmacoeconomics and Outcomes Research (ISPOR) checklist for retrospective database studies (248). The ISPOR tool is one of two, specifically designed for evaluating the quality of published studies that use real-world, health-related retrospective database evidence (249). Out of the two tools, the ISPOR checklist for retrospective databases was chosen as this enabled assessment of details of data sources, statistical results, and the paper's generalisability (248). These were all areas that are not always fully addressed with the scope of "standard" appraisal tools for observational studies, such as the Newcastle-Ottawa Scale (250). The ISPOR tool was revised and expanded to be relevant for the research question and included ten features each scored as 0, 0.5 or 1 (Appendix 2.6). Appraisal was conducted independently by two authors (RN and RD), with disagreements resolved through discussion with a third author (AT), who has consensus. No studies were weighted or excluded from either the narrative review or meta-analysis (e.g. as a sensitivity analysis) based on poor quality as identified by the quality appraisal. Some reviews address the issue of study quality through undertaking a sensitivity analysis excluding studies appraised as being low quality (251). This approach was not taken in this analysis. The rationale for this was that numerically quantifying a cut off for a "poor" study quality is challenging and subjective – especially as the criteria for appraising studies varies by the tool used (252) and consistent judgement between tools is currently lacking (253).

2.3.3 Synthesis of Evidence

A narrative synthesis was undertaken. Data was synthesised using a summary of findings table. This is reported as both a full set of papers, as well as a full set of studies (when taking into consideration the multiple publications reporting identical or heavily overlapping study populations).

Where not reported, percentages utilising novel anti-cancer therapies and/or predictive biomarker test by socio-economic sub-group were calculated from data reported in the paper or supplied by authors: unadjusted ORs for low, compared to high, SES were computed for test/therapy receipt. Studies were heterogeneous in terms of the following: outcomes (test/therapy receipt or non-receipt); socio-economic measure comparisons made; whether ORs (crude or adjusted) were reported; and the variables that any adjusted ORs were controlled for. Unadjusted ORs were therefore computed to enable inclusion in the meta-analysis of as many studies as possible in a consistent way. "Low" SES was defined as the lowest socio-economic sub-group in each article and "high" SES as the top sub-group.

A meta-analysis (statistical combination of results from several studies) was undertaken on a more limited subset of the full set of paper/studies included in the narrative synthesis. Meta-analyses were performed using random-effects, Mantel-Haenszel (M-H) methods (254). This meant that a summary statistic, in this instance an OR, was calculated in the same way for each included study to describe the dichotomous data (255). These ORs assessed the likelihood of: (i) test utilisation; and (ii) treatment utilisation by low versus high SES. The meta-analysis then combined study ORs to provide a summary of the effect estimate by calculating a weighted average of the ORs estimated for each individual study (255). As random-effects models were deployed, the analyses considered the effect of both within study variance (chance) and between study variance (heterogeneity) (255). Eligibility criteria for studies to be included in the meta-analysis were as follows: unadjusted low and high socio-economic utilisation data for one measure of SES reported and an independent sampling frame (no data overlap with another study/paper). As studies varied in their measures of SES used, and sometimes reported multiple measures of SES, a hierarchy was developed to determine which SES measure data to use in the meta-analysis. This was as follows in the primary analyses, results relating to: (i) an income measure; (ii) failing that an education measure; (iii) or otherwise, the only reported SES measure were included. This approach reflected the dominance of USA studies within the evidence-base, where there are cost implications for drug access (21). The decision was taken to not restrict meta-analyses to only studies pertaining to the same measure of SES measure (e.g. income) as measurement of these single measures varied considerably between studies and were therefore not directly comparable; moreover such a restriction could lead to the exclusion of very relevant literature. Where multiple papers included study populations from the same or related databases that overlapped in terms of period of diagnosis/treatment, the publication reporting the largest total number of patients was entered into the primary meta-analysis. A decision tree to illustrate meta-analysis study inclusion criteria is shown in Appendix 2.7.

For predictive biomarker tests, results were grouped by cancer site (breast, colorectal, lung and melanoma). Those for novel anti-cancer therapies were grouped by drug class (targeted therapy, biologic, and immunotherapy), while separate pre-specified sub-group analyses were conducted for: breast cancer; lung cancer and all other cancers (sub-grouped by cancer type: colorectal, head and neck, hepatobiliary, melanoma, mixed, and renal cell). A final *post hoc* sub-group analysis was performed for the Surveillance, Epidemiology, and End Results program (SEER)[3] versus non-SEER registry studies. Testing for sub-group differences were computed where appropriate. Two *post hoc* sensitivity analyses (one involving substituting included studies with those excluded due to overlapping sampling frames and the other exploring the United States of America (USA) versus non-USA healthcare settings) were conducted to determine the robustness of the results to the individual papers included. It was not possible to specify these analyses beforehand because the scope of the literature was uncertain at the outset given the breadth and novelty of the research topic. The decision to undertake sensitivity analyses *post hoc* in such instances is supported by Cochrane (257).

Sensitivity analyses were not undertaken on study quality as numerically quantifying what constitutes a "good study" was deemed difficult and subjective. This decision is supported by the literature (258) where differentiating study quality in the absence of a gold standard coding tool is known to be challenging (259–261) and subject to misclassification (often subjected to "low quality" classification when study results are unexpected) (262). Whilst inclusion of only high-quality appraised studies may produce a more valid meta-analysis of true effect estimates that minimised bias and validity concerns of low study quality inclusion. Notably this refers to the "garbage in, garbage out" concern as coined by Eysenck (1978) (263), statisticians acknowledge that in practice this is not always feasible (260). Traditionally, meta-analyses have therefore tended to include all potentially relevant studies regardless of quality and instead address validity concerns afterwards - this was also the approach taken in this analysis (258).

The I^2 statistic was calculated to estimate the degree of statistical heterogeneity (the percentage of variation that is not due to chance) (264). There are different ways that high heterogeneity is defined and Higgins *et al.* (2003) caution a defined categorisation, arguing for a broad I^2 representation of 0%, 25%, 50% and 75% conferring: none; low; moderate; and high heterogeneity respectively (264). The Cochrane Handbook reports a classification of 75%-100% or more as representing "considerable heterogeneity", though as they also define 50-90%

[3] SEER is a consortium of USA population-based registries (256).

as “may represent substantial heterogeneity” (265), the cut off for “high heterogeneity” seems rather arbitrary and inconsistency has caused uncertainty (266). For this review, high heterogeneity was therefore defined as ≥75%. Finally, funnel plots were produced to assess publication bias in analyses of ten plus studies (267). Statistical analyses were conducted using RevMan 5.3.

2.4 Results

2.4.1 Search Results

The search identified 17,047 citations. After removal of duplicates, titles, and abstracts of 10,722 records were screened for eligibility. After title and abstract screening, 551 records progressed to full text review. Overall, 62 papers (reporting 58 independent studies) met the inclusion criteria (Figure 2.1) and were included in the review. Of these 58 studies, 48 were included in the narrative review. The remaining 10 studies (Appendix 2.8) had no denominator populations or only reported an average measure of SES (e.g. mean household income) and were excluded from inclusion in the meta-analysis and are not discussed further (268–277). Of the narrative review, eight studies reporting utilisation data for predictive biomarker tests (278–285), 37 studies (41 papers) reporting utilisation data for novel anti-cancer therapies (286–304,304–322) and three studies reporting both (323–325). Finally, of the 48 studies included in the narrative review, only 38 of these studies meet the criteria for inclusion in the meta-analysis.

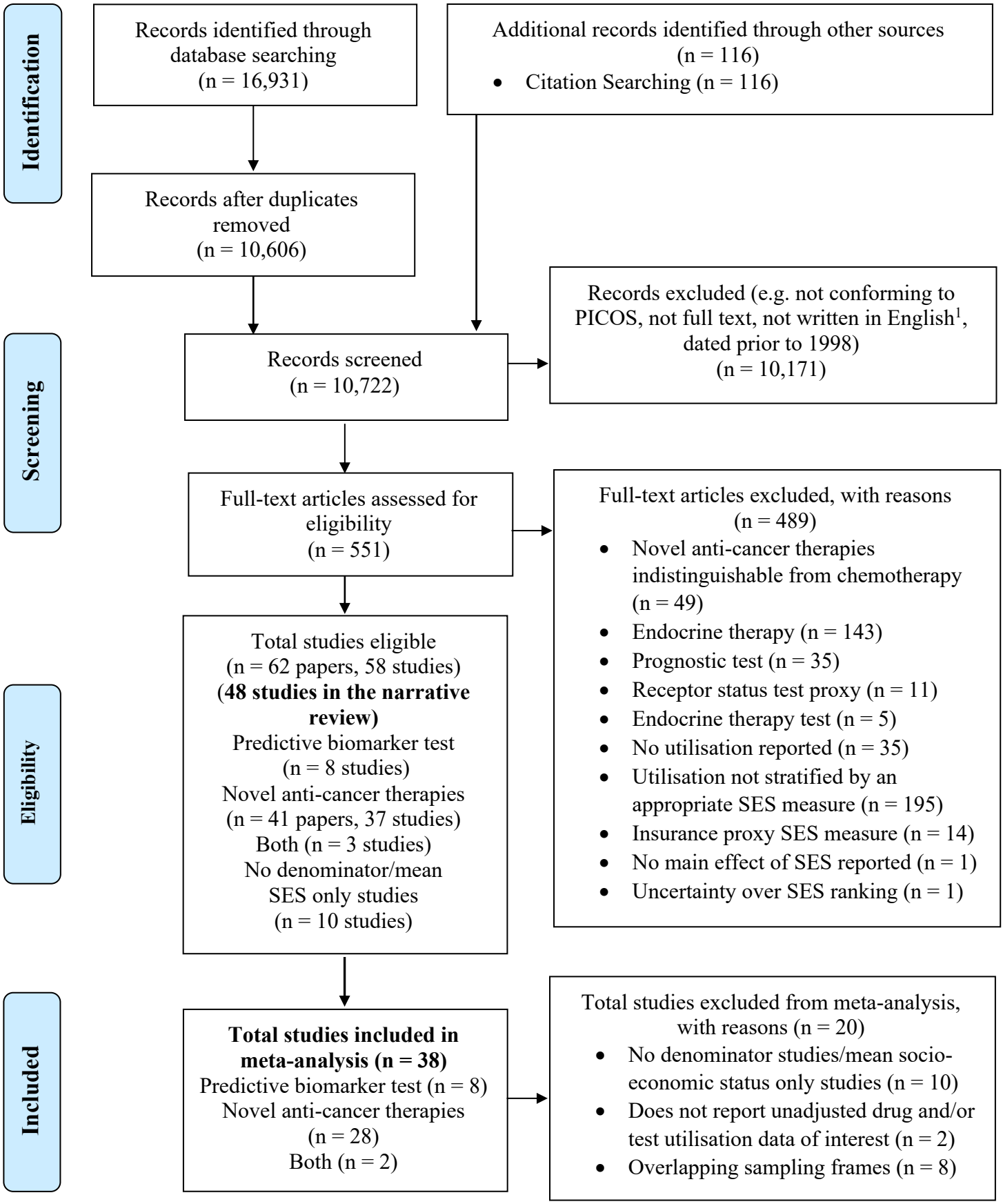


Figure 2.1 Study selection according to the PRISMA statement. SES: Socio-economic status.
[1]46 citations were not written in English (French n=8; Chinese n=8; Portuguese n=6; Spanish n=4; German n=4; Dutch n=3; Italian n=3; Danish n=3; Japanese n=2; Polish n=2; Czech n=1; Russian n=1; and Swedish n=1)

2.4.2 Study Characteristics

The 48 included studies covered: 7 cancers; 5 predictive biomarker tests; and 11 novel anti-cancer therapy classifications, of which bevacizumab (12 studies) (298–302,311–313,315,316,321,322) and trastuzumab (11 studies) (286–296) were most common. There were 6 studies on immunotherapies (297,303,305,309,310,317,318). Most studies were conducted in the USA (n = 42) (278,279,282–292,297–325), and a majority analysed SEER registry data (n =27) (278,279,282,286–290,298,299,302,304,306,307,311–316,318–321,323–325) (Appendix 2.9). Of the SEER data studies, 19 (278,279,282,286–290,298,299,302,304,306,307,311–316,318–321,323–325) were SEER Medicare (i.e. included patients ≥65). The remaining studies were from: Canada (4 studies) (280,293–295); China (1 study) (296); and Ireland (1 study) (281). Forty-six studies reported one or more area-based SES measures, and only two studies utilised individual-based measures (patient reported educational attainment, employment and eligibility for low income subsidy for Medicare part D) (291,325). Six SES measures were reported: poverty (9 studies with an area-based measure (279,283,286,288,306,313,314,320,325) and 1 study where the unit of measure was unreported (304)); income (30 studies with an area-based measure (278,280,282,284,285,287,289,290,292–300,303,305,307–309,311,312,317–319,323,324) and one study with an individual-based measure (325)); education (20 area-based studies (283,284,287,289,290,296–301,303,308,309,311,317,318,322–324) and one individual-based measure study (291)); individual-based employment measure (1 study) (291); area-based deprivation measures (1 study) (281); and area-based SES aggregate score measures (4 studies) (302,315,316,321). For nine studies, utilisation by SES group was only available as percentages (281,286,289,309,311,313,318,323,324). Study characteristics are summarised in Table 2.1.

Seven papers, pertaining to four studies, reported the same data from the same registry (California Cancer Registry, (CCR) Cancer Research Network (CRN) and the Ontario Cancer Registry (OCR)) (295,300,302,326–329). Sixteen papers (covering 8 studies) overlapped in their study populations (cancer site, stage, years of diagnosis time frames, patients' age) (286–289,293,294,298–301,306,307,311,312,324,325). Two studies did not report unadjusted drug and/or test utilisation data (297,315). This left 38 studies (equating to 1,036,125 patients) which were included in the meta-analysis (278–286,288,290–292,294–296,299,301–306,308–311,313,314,316–323,325).

Table 2.1 Characteristics of included studies

	Sampling Frame					SES		Utilisation by SES Grouping (Number, %)						
Study	**Country**	**Data Source**	**Study Population[a]**	**Predictive Biomarker Test/Biological and Precision Medicine Overall Utilisation (Number, %)**	**Comparator**	**Unit**	**Measure**	**Lowest SES Group**				**Highest SES Group**	**Utilisation in Low SES (OR & 95% CI)[b]**	**QA**
Predictive Biomarker Testing														
Pensa *et al.* (2009)* (278)	USA	Connecticut Tumor Registry - SEER Database	Connecticut 2000 - 2003 Breast Cancer n = 1,364	HER2 Test n = 894 (65.5)	No HER2 Test	Census	Median Household Income of Town	Lowest Tertile (Low) 299/462 (64.7) OR: 0.94 (0.75 - 1.20)	Middle /Highest Tertile (High) 595/902 (66.0) OR: Ref				0.95 (0.74 - 1.21)	5.5
Lund *et al.* (2010)* (279)	USA	Atlanta SEER Registry & Georgia Comprehensive Cancer Registry	Fulton and Dekalb Counties in Metropolitan Atlanta 2003 - 2004 Stage I - IV Breast Cancer n = 1,842	HER2 Test n = 1,660 (90.1)	No HER2 Assay	Census Tract	% Living Below the Federally Defined Poverty Line	≥ 20% (Low) 286/319 (89.7) P Value: 0.9565	10 < 20% 390/433 (90.1)	5 < 10% 435/479 (90.8)	< 5% (High) 548/607 (90.3)		0.93 (0.58 - 1.51)	6.5
Ferrusi *et al.* (2013)* (280)	Canada	OCR	Ontario 01/2006 - 12/2007 Early Stage (I - III) Breast Cancer n = 13,396	Documented HER2 Test n = 8,854 (66.1)	Undocumented HER2 Test	Census Tract (Postcode)	Household Income	Q1 (Low) 1,563/2,334 (67.0) Q1, 2, 3 or 4 vs Q5 OR: 1.00 (0.85 - 1.17) P Value: Not significant at 0.05	Q2 1,667/2,577 (64.7)	Q3 1,770/2,649 (66.8)	Q4 1,856/2,812 (66.0)	Q5 (High) 1,977/2,991 (66.1)	1.04 (0.93 - 1.17)	8.5

Table 2.1 Continued

	Sampling Frame					SES		Utilisation by SES Grouping (Number, %)						
Study	**Country**	**Data Source**	**Study Population[a]**	**Predictive Biomarker Test/Biological and Precision Medicine Overall Utilisation (Number, %)**	**Comparator**	**Unit**	**Measure**	**Lowest SES Group**				**Highest SES Group**	**Utilisation in Low SES (OR & 95% CI)[b]**	**QA**
de Camargo Cancela *et al.* (2015) *(281)	Ireland	National Cancer Registry Ireland	Ireland 2006 - 2008 Stage I - IV Breast Cancer n = 7,619[c]	HER2 Test n = 6,529 (85.7)	ER Test/ PR Test/ Any Hormone Receptor Tests	Area of Residence	Deprivation Status	Q1 (Low)	Q2	Q3	Q4	Q5 (High)	0.84 (0.69 - 1.02)	7
								1,697/1,990 (85.3)[c]	970/1,135 (85.5)[c]	827/986 (83.9)[c]	904/1,042 (86.8)[c]	1,476/1,691 (87.3)[c]		
								IRR (No HER2 Test): 1.15 (0.98 - 1.36)	IRR (No HER2 Test): 1.14 (0.95 - 1.38)	IRR (No HER2 Test): 1.27 (1.05 - 1.53)	IRR (No HER2 Test): 1.04 (0.85 - 1.27)	IRR (No HER2 Test): Ref		
Greenbaum *et al.* (2017)*. (282)	USA	New Mexico Tumour Registry Participating in SEER	New Mexico Residents 2010 - 2013 Stage IV Colorectal Cancer n = 637	KRAS Test n = 245 (38.5)	NR	Census Tract	Annual Income per capita	$5,051 - $15,656 (Low)	$15,662 - $23,034	$23,126 - $32,042	$32,138 – $84,620 (High)		0.77 (0.47 - 1.26)	8.5
								52/153 (34.0)	62/153 (40.5)	60/153 (39.2)	61/152 (40.1)			
								OR: Ref	OR: 1.48 (0.88 - 2.48)	OR: 1.27 (0.68 - 2.35)	OR: 1.21 (0.63 - 2.34)			
								P Value: 0.618						

Table 2.1 Continued

Sampling Frame						SES		Utilisation by SES Grouping (Number, %)			
Study	**Country**	**Data Source**	**Study Population[a]**	**Predictive Biomarker Test/Biological and Precision Medicine Overall Utilisation (Number, %)**	**Comparator**	**Unit**	**Measure**	**Lowest SES Group**	**Highest SES Group**	**Utilisation in Low SES (OR & 95% CI)[b]**	**QA**
Rico *et al.* (2016)* (283)	USA	10 Centers for Disease Control & Prevention National Program of Central Cancer Registries	Alaska, California (13 Counties in Sacramento area), Colorado, Florida (5 Miami Metro Counties), Idaho, Louisiana, New Hampshire, North Carolina, Rhode Island & Texas. 2011 Stage: Metastatic Colorectal Cancer n = 3,608	KRAS Test n = 992 (27.5)	Not Tested	Census Tract	% People Living Under the Federal Poverty Level	≥ 20% (Low) 251/968 (25.9) OR: 1.13 (0.91 - 1.40) P Value: 0.2665	< 20% (High) 733/2,619 (28.0) OR: Ref P Value: 0.2201	0.90 (0.76 - 1.07)	9.5
						Census Tract	% Without High School Education	≥ 25% (Low) 212/902 (23.5) OR: 0.86 (0.69 - 1.07) P Value: 0.1828	< 20% (High) 773/2,689 (28.7) P Value: **0.0023** OR: Ref	0.76 (0.64 - 0.91)	

Table 2.1 Continued

Sampling Frame						SES		Utilisation by SES Grouping (Number, %)					
Study	**Country**	**Data Source**	**Study Population[a]**	**Predictive Biomarker Test/Biological and Precision Medicine Overall Utilisation (Number, %)**	**Comparator**	**Unit**	**Measure**	**Lowest SES Group**			**Highest SES Group**	**Utilisation in Low SES (OR & 95% CI)[b]**	**QA**
Webster *et al.* (2013)* (284)	USA	Virtual Data Warehouse Tumour Registry Files	7 CRN Sites Across the USA (Kaiser Permanente, Henry Ford Health System & Health Partners) 01/2004 - 12/2009 Stage III (Progressed to Distant Metastatic Disease) & IV Colorectal Cancer n = 1,188	KRAS Tested n = 428 (36.0)	Not KRAS Tested	Census Tract	Median Household Income	< $40K (Low)	$40K - $59K	$60K - $79K	≥ $80K (High)	0.73 (0.48 - 1.10)	10
								95/316 (30.1)	142/405 (35.1)	90/231 (39.0)	63/170 (37.1)		
								P Value: **0.0024**					
								OR: 0.80 (0.50 - 1.40)	OR: Ref	OR: 1.30 (1.10 - 1.40)	OR: 1.20 (1.10 - 1.50)		
								P Value: 0.4753					
						Census Tract	% with a High School Education	< 50% (Low)	50 - 69%	70 - 89%	≥ 90% (High)	0.52 (0.12 - 1.76)	
								4/16 (25.0)	36/99 (36.4)	197/580 (34.0)	166/426 (39.0)		
								P Value: **< 0.0001**					
								OR: 0.70 (0.10 - 3.30)	OR: 1.4 (1.00 - 1.80)	OR: Ref	OR: 1.20 (1.00 - 1.40)		
								P Value: 0.9924					
Enewold *et al.* (2017)*[d] (323)	USA	Hospital Medical Records, Contact with Treating Physicians and Others Involved Individual's Care	NCI POC Study of SEER Patients (28% of US Population) 2011 Stage: Metastatic Melanoma n = 520[c]	BRAF Test n = 242 (46.5)	NR	Census Tract	Median Household Income	$15,769 - $41,350 (Low)	$41,351 - $55,155	$55,156 - $73,178	$73,179 - $163,393 (High)	0.80 (0.48 - 1.35)	4.5
								52/124 (41.9)[c]	61/128 (47.7)[c]	64/130 (49.2)[c]	63/133 (47.4)[c]		
								P Value: 0.70					
						Census Tract	% of Individuals Aged 25+ with at Least a High School Education	28.30% - 77.28% (Low)	77.39% - 86.14%	86.15% - 91.26%	91.27% - 100.00% (High)	0.66 (0.39 - 1.11)	
								50/126 (40.0)[c]	60/126 (47.6)[c]	65/133 (49.0)[c]	65/130 (50.0)[c]		
								P Value: 0.46					

Table 2.1 Continued

Sampling Frame						SES		Utilisation by SES Grouping (Number, %)				
Study	**Country**	**Data Source**	**Study Population[a]**	**Predictive Biomarker Test/Biological and Precision Medicine Overall Utilisation (Number, %)**	**Comparator**	**Unit**	**Measure**	**Lowest SES Group**		**Highest SES Group**	**Utilisation in Low SES (OR & 95% CI)[b]**	**QA**
Enewold & Thomas (2016)*[d] (324)	USA	Medical Records and Treating Physicians Contacted	NCI POC Study from SEER 2010 Stage IV NSCLC n = 764[c]	EGFR Test n = 152 (19.9)	NR	Census Tract	Median Income	< $43,000 (Low) 49/296 (16.5)[c] P Value: 0.39	$43,000 - $62,000 46/248 (18.5)[c]	>$62,000 (High) 55/220 (25.0)[c]	0.60 (0.38 - 0.94)	5
Palazzo *et al.* (2019)*[d] (325)	USA	SEER - Medicare	17 SEER Registries Comprising Approx 28% of US Population 2007 - 2011 Age ≥ 65 Stage IV NSCLC n = 9,900	Genetic Test n = 1,040 (10.5)	No Genetic Test	Eligibility for or Receipt of Low Income Subsidy for Medicare Part D	Income	Low Income (Low) 319/4,212 (7.6) P Values: **< 0.0001** OR: 0.78 (0.57 - 1.08) P Values: 0.1335	Not Low Income (High) 721/5,688 (12.7) OR: Ref		0.56 (0.49 - 0.65)	10
						Census Tract	High Poverty Location	High Poverty (Low) 439/5,309 (8.3) P Value: **< 0.0001** OR: 0.82 (0.62 - 1.08) P Value: 0.1553	Not High Poverty (High) 554/4,202 (13.2) OR: Ref		0.59 (0.52 - 0.68)	

Table 2.1 Continued

Sampling Frame						SES		Utilisation by SES Grouping (Number, %)						
Study	**Country**	**Data Source**	**Study Population[a]**	**Predictive Biomarker Test/Biological and Precision Medicine Overall Utilisation (Number, %)**	**Comparator**	**Unit**	**Measure**	**Lowest SES Group**				**Highest SES Group**	**Utilisation in Low SES (OR & 95% CI)[b]**	**QA**
Presley *et al.* (2018)* (285)	USA	Flatiron Health Database	250 Cancer Clinics with 1.5 Million Active Patients (Received Care at 1 of 191 Oncology Practices) 01/2011 - 07/2016 Stage IIIB - IV Non-squamous NSCLC n = 5,688	Routine Test (EGFR and/or ALK) n = 4,813 (84.6)	Broad Based Genomic Sequencing (Multigene Panel Testing More than 30 Genes)	Zip Code	Median Household Income	Q1 (Low) 484/555 (87.2) P Value: **< 0.001**	Q2 637/732 (87.0)	Q3 927/1,090 (85.0)	Q4 1,037/1,214 (85.4)	Q5 (High) 1,612/1,969 (81.9)	1.51 (1.14 - 2.01)	6
Breast Cancer: Novel Anti-Cancer Therapies														
Du *et al.* (2011)* (286)	USA	SEER - Medicare	16 USA Cancer Registries 1998 - 2005 Age ≥ 65 Stage I - IV & Unstaged n = 47,806[c]	Trastuzumab & Anthracycline n = 460 (1.0) Trastuzumab & No Anthracycline n = 414 (0.9)	Anthracycline/ Other Chemotherapy/ No Chemotherapy	Census Tract	% of Persons Living Below the Poverty Line	≥ 12% (Low) 221/11,918 (1.9)[c]	6.63% - 11.99% 198/11,775 (1.7)[c]	3.63% - 6.62% 230/11,728 (2.0)[c]	≤3.62% (High) 222/11,855 (1.9)[c]		0.99 (0.82 - 1.20)	5
Vaz-Luis *et al.* (2015) (287)	USA	SEER - Medicare	28% of US Population 10/1998 - 12/2009 Age ≥ 66 Stage IV White Women N = 3,748	HER2+ (Initiated Trastuzumab) n = 347 (9.3)	HER2- (No Trastuzumab)	Census Tract	Median Household Income	Q1 (Low) 64/742 (8.6)[c]	Q2 78/944 (8.3)[c]	Q3 96/1,015 (9.5)[c]	Q4 (High) 109/1,047 (10.4)[c]		0.81 (0.58 - 1.14)	5
						Census Tract	% With High School Diplomas	Q1 (Low) 65/788 (8.2)[c]	Q2 87/931 (9.3)[c]	Q3 99/983 (10.1)[c]	Q4 (High) 96/1,046 (9.2)[c]		0.89 (0.63 - 1.25)	

Table 2.1 Continued

Sampling Frame						SES		Utilisation by SES Grouping (Number, %)					
Study	**Country**	**Data Source**	**Study Population[a]**	**Predictive Biomarker Test/Biological and Precision Medicine Overall Utilisation (Number, %)**	**Comparator**	**Unit**	**Measure**	**Lowest SES Group**			**Highest SES Group**	**Utilisation in Low SES (OR & 95% CI)[b]**	**QA**
Reeder-Hayes *et al.* (2016)* (288)	USA	SEER - Medicare	28% of US Population 2010 - 2011 Age ≥ 66 Stage I - III n = 1,362	Trastuzumab Treated Users n = 672 (49.3)	Untreated (No Trastuzumab)	Census Tract	Residents Below the Poverty Level	≥ 20% (Low)	10 - 19.99%	5 - 9.99%	< 5% (High)	0.71 (0.50 - 1.00)	9.5
								122/261 (46.7)	212/439 (48.3)	178/373 (47.7)	160/289 (55.4)		
								P Value: 0.1397					
								RR: 0.88 (0.76 - 1.02)	RR: 0.87 (0.77 - 0.99)	RR: 0.85 (0.75 - 0.97)	RR: Ref		
								P Value: 0.0965	P Value: **0.0377**	P Value: **0.0155**			
Vaz- Luis *et al.* (2016) (289)	USA	SEER - Medicare	28% of US Population 2010 - 2011 Age ≥ 66 Stage ≥ Ib n = 770[c]	Trastuzumab n = 428 (55.6)	No Trastuzumab	Census Tract	Median Household Income	Q1 (Low)	Q2	Q3	Q4 (High)	0.75 (0.49 - 1.14)	7.5
								103/196 (52.6)[c]	102/192 (53.1)[c]	105/184 (57.1)[c]	118/198 (59.6)[c]		
								P Value: 0.45					
								OR (Non-Receipt): Ref	OR (Non-Receipt): 1.22 (0.78 - 1.89)	OR (Non-Receipt):1.40 (0.82 - 2.20)	OR (Non-Receipt):1.60 (0.83 - 3.10)		
								P Value: 0.56					
						Census Tract	High School Diploma Rate	Q1 (Low)	Q2	Q3	Q4 (High)	0.55 (0.36 - 0.83)	
								98/204 (48.0)[c]	111/188 (59.0)[c]	99/187 (52.9)[c]	120/191 (62.8)[c]		
								P Value: **0.02**					
								OR (Non-Receipt): Ref	OR (Non-Receipt): 2.32 (1.24 - 4.35)	OR (Non-Receipt): 1.21 (0.78 - 1.91)	OR (Non-Receipt): 1.68 (1.11 - 2.53)		
								P Value: **0.05**					

Table 2.1 Continued

Sampling Frame						SES		Utilisation by SES Grouping (Number, %)						
Study	**Country**	**Data Source**	**Study Population[a]**	**Predictive Biomarker Test/Biological and Precision Medicine Overall Utilisation (Number, %)**	**Comparator**	**Unit**	**Measure**	**Lowest SES Group**				**Highest SES Group**	**Utilisation in Low SES (OR & 95% CI)[b]**	**QA**
Tsai *et al.* (2017)* (290)	USA	3 SEER Registries, NYSCR, OCISS Linked to Healthcore Inc Claims	5 US States 01/ 2006 - 12/2011 Age < 64 Stage I - III HER2+ n = 934	Trastuzumab n = 680 (72.8)	NR	Census Tract	Median Household Income	Q1 (Low)	Q2	Q3	Q4	Q5 (High)	1.14 (0.69 - 1.90)	6
								104/140 (74.3)	121/176 (68.8)	139/185 (75.1)	151/204 (74.0)	155/216 (71.8)		
						Census Tract	% Completing College	< 20%	20% to < 30%	30% to < 40%	≥ 40%		1.07 (0.60 - 1.89)	
								127/169 (75.1) OR: Ref	289/406 (71.2) OR: 0.82 (0.50 - 1.35)	166/227 (73.1) OR: 0.84 (0.48 - 1.45)	88/119 (73.9) OR: 0.95 (0.51 - 1.79)			
									P Value: 0.44	P Value: 0.53	P Value: 0.89	Overall P Value: 0.85		
Freedman *et al.*(2013)* (291)	USA	NCCN Breast Cancer Outcomes Database	8 US Centers 09/2005 - 12/2008 Stage I - III n = 1,109	Trastuzumab n = 925 (83.0)	NR	Patient Reported	Educational Attainment	< High School (Low)	High School Degree	Some College	College/ Graduate Degree (High)		1.06 (0.47 - 2.73)	8
								46/54 (85.2) P Value: 0.911	131/161 (81.4)	182/218 (83.5)	341/404 (84.4)			
								OR 1.58 (0.98 - 2.52)	OR: 0.84 (0.58 - 1.23)	OR: 0.94 (0.67 - 1.32)	OR: Ref			
						Patient Reported	Employment	Unemployed (Low)	Retired	Homemaker	Employed/ Student (High)		1.34 (0.61 - 3.35)	
								61/69 (88.4) P Value: **0.002**	110/153 (71.9)	148/178 (83.1)	491/577 (85.1)			
								OR = 1.11 (0.56 - 2.21)	OR: 1.37 (0.96 - 1.94)	OR: 0.89 (0.69 - 1.15)	OR: Ref			

Table 2.1 Continued

	Sampling Frame					SES		Utilisation by SES Grouping (Number, %)						
Study	**Country**	**Data Source**	**Study Population[a]**	**Predictive Biomarker Test/Biological and Precision Medicine Overall Utilisation (Number, %)**	**Comparator**	**Unit**	**Measure**	**Lowest SES Group**			**Highest SES Group**		**Utilisation in Low SES (OR & 95% CI)[b]**	**QA**
Haas *et al.* (2011)* (292)	USA	Aetna Medical Records	US Women Receiving Aetna Health Coverage 07/2006 - 06/2007 Age: 35 - 65 Stage I - III HER2+ Test Result n = 137	Trastuzumab n = 79 (57.7)	NR	Small Area Estimation	Annual Household Income	< $40,000 (Low)	$40,000 - $74,999	$75,000 - $124,999	≥$125,000 (High)		4.21 (1.14 - 16.04)	8.5
								27/34 (79.4)	22/47 (46.8)	16/27 (59.3)	11/23 (47.8)			
								P Value: **0.02**						
								OR: 4.43 (1.22 - 16.04)	OR: 1.01 (0.34 - 3.00)	OR: 1.61 (0.47 - 5.51)	OR: Ref			
Goldhar *et al.* (2016) (293)	Canada	OCR Linked to RPBD, CIHI-DAD, OHIP, ODB, NDFP & NACRS	Ontario 2003 - 2009 Stage I - III n = 19,074	Trastuzumab and Chemotherapy n = 3,371 (17.7)	Chemotherapy Alone	Postal Code	Average Household Income	Q1 (Low)	Q2	Q3	Q4	Q5 (High)	1.06 (0.93 - 1.19)	6
								564/3,029 (18.6)	625/3,541 (17.7)	667/3,863 (17.3)	722/4,191 (17.2)	781/4,385 (17.8)		
								P Value: 0.57						
Kumachev *et al.* (2016)* (294)	Canada	OCR Linked to RPBD, CIHI-DAD, OHIP, NDFP, NACRS & ODB	Ontario 01/2004 - 12/2009 Age > 18 Stage I - III n = 33,056	Trastuzumab n = 3,391 (10.3)	NR	Postal Code	Average Neighborhood Income	Q1 (Low)	Q2	Q3	Q4	Q5 (High)	0.93 (0.83 - 1.04)	4.5
								577/5,890 (9.8)	645/6,448 (10.0)	662/6,362 (10.4)	725/6,906 (10.5)	782/7,450 (10.5)		
								P Value: 0.62						

Table 2.1 Continued

Sampling Frame						SES		Utilisation by SES Grouping (Number, %)						
Study	**Country**	**Data Source**	**Study Population[a]**	**Predictive Biomarker Test/Biological and Precision Medicine Overall Utilisation (Number, %)**	**Comparator**	**Unit**	**Measure**	**Lowest SES Group**				**Highest SES Group**	**Utilisation in Low SES (OR & 95% CI)[b]**	**QA**
Thaven - diranathan et al.* (2016 & 2018) (295,327)	Canada	OCR Linked to RPBD, CIHI-DAD, OHIP, NACRS, NDFP, CALR & ICES	14 Ontario Cancer Centers 07/2007 - 12/2012 Age > 18 Stage I - III n = 18,540	Trastuzumab Without Anthracycline/ n = 832 (4.5) Sequential Therapy (Anthracycline followed by Trastuzumab) n = 3,250 (17.5)	Anthracycline Without Trastuzumab/ Other Chemotherapy	Postal Code	Median Income	Q1 (Low) 658/3,088 (21.3)	Q2 775/3,454 (22.4)	Q3 812/3,658 (22.2)	Q4 952/4,102 (23.2)	Q5 (High) 885/4,238 (20.9)	1.03 (0.91 - 1.15)	5
Li *et al.* (2018)* (296)	China	Hospital Records, Telephone Conversations & Other Means	155 Hospitals (29 Chinese Provinces) 07/2013 - 06/2014 Age ≥ 18 Stage I - III n = 4,994	Trastuzumab n = 1,487 (29.8)	Non-Trastuzumab	NR	Household Income	< 10,000 Yuan (Low) 12/54 (22.2) P Value: **< 0.001**	10,000 - 30,000 Yuan 55/276 (19.9)	30,000 - 50,000 Yuan 82/284 (28.9)	> 50,000 Yuan (High) 168/357(47.1)		0.32 (0.15 - 0.65)	6
						NR	Education Level	Primary School or Lower (Low) 112/469 (23.9) P Value: **< 0.001**	High School 238/738 (32.2)	College or Higher (High) 133/266 (50.0)			0.31 (0.22 - 0.44)	

Table 2.1 Continued

Sampling Frame						SES		Utilisation by SES Grouping (Number, %)						
Study	**Country**	**Data Source**	**Study Population[a]**	**Predictive Biomarker Test/Biological and Precision Medicine Overall Utilisation (Number, %)**	**Comparator**	**Unit**	**Measure**	**Lowest SES Group**			**Highest SES Group**		**Utilisation in Low SES (OR & 95% CI)[b]**	**QA**
Shih *et al.* (2009)[f] (297)	USA	NCDB	Captures 75% of Newly Diagnosed Cancer Cases 1998 - 2004 Stage: IV n = 42,804	Immunotherapy n = 1,723 (4.0)	NR	Zip Code	% Without a High School Degree	29%+ OR: 0.98 (0.67 - 1.44)[g]	20% - 28.9% OR: 0.77 (0.55 - 1.08)[g]	14% - 19.9% OR: 0.92 (0.68 - 1.24)[g]	< 14% OR: Ref[g]		[h]	7
						Zip Code	Median Household Income	< $30,000 OR: Ref[g]	$30,000 - $34,999 OR: 1.42 (0.99 - 2.06)[g]	$35,000 - $45,999 OR: 1.11 (0.77 - 1.62)[g]	$46,000+ OR:1.05 (0.71 - 1.55)[g]			
Non-Small-Cell Lung Cancer: Novel Anti-Cancer Therapies														
Zhu *et al.* (2012) (298)	USA	SEER - Medicare	17 US Registries (28% of US Population) 2006 - 2007 Age ≥ 65 Stage IIB - IV Non-squamous n = 1,500 n (PSM) = 636	Bevacizumab with Carboplatin and Paclitaxel n = 318 (21.2) n (PSM) = 318 (50.0)	Carboplatin & Paclitaxel	Census Tract	Median Household Income	Q1 (Low) 60/300 (20.0) P Value: 0.20 PSM: 60/123 (48.8) P Value: 0.88	Q2 58/300 (19.3) PSM: 58/106 (54.7)	Q3 63/300 (21.0) PSM: 63/128 (49.2)	Q4 58/300 (19.3) PSM: 58/116 (50.0)	Q5 (High) 79/300 (26.3) PSM: 79/163 (48.5)	0.70 (0.47 - 1.04)	8
						Census Tract	% of Persons Older than Age 25 with Some College Education	Q1 (Low) 66/300 (22.0) P Value: 0.38 PSM: 66/147 (44.9) P Value: 0.33	Q2 58/302 (19.2) PSM: 58/104 (55.8)	Q3 62/300 (20.7) PSM: 62/117 (53.0)	Q4 57/298 (19.1) PSM: 57/125 (45.6)	Q5 (High) 75/300 (25.0) PSM: 75/143 (52.4)	0.85 (0.57 - 1.26)	

Table 2.1 Continued

Sampling Frame						SES		Utilisation by SES Grouping (Number, %)						
Study	**Country**	**Data Source**	**Study Population[a]**	**Predictive Biomarker Test/Biological and Precision Medicine Overall Utilisation (Number, %)**	**Comparator**	**Unit**	**Measure**	**Lowest SES Group**				**Highest SES Group**	**Utilisation in Low SES (OR & 95% CI)[b]**	**QA**
Langer *et al.* (2014)* (299)	USA	SEER - Medicare	20 Geographic Areas (29% of US Population) 2006 - 2009 Age ≥ 65 Stage IIIB - IV Non-squamous n = 1,706	Bevacizumab with Carboplatin & Paclitaxel n = 592 (34.7)	Carboplatin & Paclitaxel	Census Tract	Median Household Income	Q1(Low) 96/340 (28.2) P Value: **0.039**	Q2 123/341 (36.1)	Q3 129/343 (37.6)	Q4 113/340 (33.2)	Q5 (High) 131/342 (38.3)	0.63 (0.45 - 0.88)	6
						Census Tract	% of Persons Older than 25 Years with Some College Education	Q1 (Low) 117/340 (34.4) P Value: 0.317	Q2 111/341 (32.6)	Q3 111/343 (32.4)	Q4 119/341 (34.9)	Q5 (High) 134/341 (39.3)	0.81 (0.59 - 1.12)	
Ritzwoller et al. (2014), Delate *et al*. (2014) & Carroll *et al.* (2015) (300,328,329)	USA	CRN's Virtual Data Warehouse	4 US HMOs 2005 - 2010 Age ≥ 21 Stage IIIB - IV Non-squamous n = 1,109	Bevacizumab with Carboplatin & Paclitaxel n = 198 (17.9)	Carboplatin & Paclitaxel	Census Tract	% College Educated	Q1 (Low) 30/224 (13.4) P Value: 0.17	Q2 50/226 (22.1)	Q3 38/224 (17.0)	Q4 42/217 (19.4)	Q5 (High) 38/218 (17.4)	0.73 (0.42 - 1.27)	6.5
						Census Tract	Median Family Income	Q1 (Low) 36/219 (16.4) P Value: 0.47	Q2 44/228 (19.3)	Q3 47/221 (21.3)	Q4 38/224 (17.0)	Q5 (High) 33/217 (15.2)	1.10 (0.63 - 1.90)	
Menter *et al.* (2016)* (301)	USA	CRN's Virtual Data Warehouse	4 Kaiser Permanente Regions 01/2005 - 12/2011 Age ≥ 21 Stage IIIB - IV Non-squamous n = 1,813 n (PSM) = 632	Bevacizumab with Carboplatin & Paclitaxel n = 348 (19.2) n (PSM) = 122 (19.3)	Carboplatin & Paclitaxel	Census	Education	Rank 1 (Low) 66/329 (20.1) PSM: 14/97 (14.4)	Rank 2 51/338 (15.1) PSM: 33/122 (27.0)	Rank 3 76/374 (20.3) PSM: 23/115 (20.0)	Rank 4 67/382 (17.5) PSM: 28/166 (16.9)	Rank 5 (High) 88/390 (22.6) PSM: 24/114 (21.1)	0.86 (0.59 - 1.25) 0.63 (0.28 - 1.38)	5.5

Table 2.1 Continued

Sampling Frame						SES		Utilisation by SES Grouping (Number, %)						
Study	**Country**	**Data Source**	**Study Population[a]**	**Predictive Biomarker Test/Biological and Precision Medicine Overall Utilisation (Number, %)**	**Comparator**	**Unit**	**Measure**	**Lowest SES Group**				**Highest SES Group**	**Utilisation in Low SES (OR & 95% CI)[b]**	**QA**
Maguire *et al.* (2019a) & ***Maguire et al. (2019b)**** [i](302,326)	USA	CCR Affiliated with SEER	California, USA 2012 - 2014 Age ≥ 20 Stage IV n = 17,254	Bevacizumab Based (Alone or in Combination with Chemotherapy) n = 530 (3.1) Pemetrexed & Bevacizumab Based (Together or with a Platinum Agent) n = 635 (3.7)	Platinum Doublets/ Pemetrexed Based/ Single Agents/ TKIs/ Chemotherapy/ No Treatment/ Unknown	Census Block	SES (Aggregate Measure of Education, Occupation, Unemployment, Household Income, Poverty, Rent & Home Price)	Q1 (Low) 152/2,888 (5.3) Bevacizumab Based OR: 0.60 (0.43 - 0.85) Pemetrexed & Bevacizumab Based OR: 0.40 (0.29 - 0.54)	Q2 204/3,530 (5.8) Bevacizumab Based OR: 0.71 (0.52 - 0.97) Pemetrexed & Bevacizumab Based OR: 0.47 (0.36 - 0.62)	Q3 237/3,703 (6.4) Bevacizumab Based OR: 0.82 (0.61 - 1.09) Pemetrexed & Bevacizumab Based OR: 0.50 (0.39 - 0.65)	Q4 272/3,771 (7.2) Bevacizumab Based OR: 0.93 (0.70 - 1.23) Pemetrexed & Bevacizumab Based OR: 0.62 (0.48 - 0.79)	Q5 (High) 300/3,362 (8.9)	0.57 (0.46 - 0.70)	8.5
				TKIs [i] n = 1,711 (9.9)	Platinum Doublets/ Pemetrexed Based/ Bevacizumab Based/ Pemetrexed & Bevacizumab/ Single Agents/ Chemotherapy/ No Treatment/ Unknown	Census Block	SES (Aggregate Measure of Education, Occupation, Unemployment, Household Income, Poverty, Rent & Home Price)	Q1 (Low) 159/2,888 (5.5) OR: 0.30 (0.24 - 0.37)	Q2 287/3,530 (8.1) OR: 0.51 (0.42 - 0.62)	Q3 340/3,703 (9.2) OR: 0.53 (0.44 - 0.63)	Q4 412/3,771 (10.9) OR: 0.66 (0.55 - 0.79)	Q5 (High) 513/3,362 (15.3) OR: Ref	0.32 (0.27 - 0.39)	

Table 2.1 Continued

Sampling Frame						SES		Utilisation by SES Grouping (Number, %)				
Study	**Country**	**Data Source**	**Study Population[a]**	**Predictive Biomarker Test/Biological and Precision Medicine Overall Utilisation (Number, %)**	**Comparator**	**Unit**	**Measure**	**Lowest SES Group**		**Highest SES Group**	**Utilisation in Low SES (OR & 95% CI)[b]**	**QA**
Palazzo e*t al*. (2019)* [d](325)	USA	SEER - Medicare	17 US Registries (28% of US Population) 2007 - 2011 Age ≥ 65 Stage IV Had Genetic Test n = 1,040	Erlotinib & had Genetic Test n = 250 (24.0)	NR	Eligibility for or Receipt of Low Income Subsidy for Medicare Part D	Income Level	Low Income (Low) 73/319 (22.9) P Value: **0.0131** OR: 0.32 (0.13 - 0.79) P Value: **0.0131**	Not Low Income (High) 177/721 (24.5) OR: Ref		0.91 (0.66 - 1.23)	9
						Census Tract	Residence in a High Poverty Location	High Poverty (Low) 102/439 (23.2) P Value: **0.0002** OR: 1.12 (0.58 - 2.17) P Value: 0.7304	Not High Poverty (High) 136/554 (24.5) OR: Ref		0.93 (0.69 - 1.26)	
Enewold & Thomas (2016)[d] (324)	USA	SEER Medical Records & Querying Treating Physicians	NCI POC Study of SEER Patients 2010 Age ≥ 20 Stage IV n = 764[c]	Erlotinib n = 70 (9.2)	NR	Census Tract	Median Income	< $43,000 (Low) 18/296 (6.0)[c] P Value: 0.11	$43,000 - $62,000 17/248 (7.0)[c]	> $62,000 (High) 32/220 (14.7)[c]	0.38 (0.20 - 0.72)	6
						Census Tract	% With a High School Education	< 77% (Low) 29/212 (13.9)[c] P Value: 0.15	77 - 89% 13/221 (5.9)[c]	> 89% (High) 24/331 (7.1)[c]	2.03 (1.10 - 3.75)	

Table 2.1 Continued

Sampling Frame						SES		Utilisation by SES Grouping (Number, %)					
Study	**Country**	**Data Source**	**Study Population[a]**	**Predictive Biomarker Test/Biological and Precision Medicine Overall Utilisation (Number, %)**	**Comparator**	**Unit**	**Measure**	**Lowest SES Group**			**Highest SES Group**	**Utilisation in Low SES (OR & 95% CI)[b]**	**QA**
Verma *et al.* (2019)*(303)	USA	NCDB	70% of US Malignancies Annually 2004 - 2015 Age ≥ 18 Stage IV n = 504,447	Immunotherapy n = 11,420 (2.3)	No Immunotherapy	NR	Income (US $/Year)	< $63,000 (Low) 7,886/360,070 (2.2) P Value: **< 0.001** OR: Ref P Value: 0.795	≥ $63,000 (High) 3,467/134,988 (2.6) OR: 0.993 (0.945 -1.044)			0.85 (0.82 - 0.88)	8
						Zip Code	Education (% With a High School Diploma)	< 80% (Low) 4,339/215,197 (2.0) P Value: **< 0.001** OR: Ref P Value: **< 0.001**	≥ 80% (High) 6,734/270,482 (2.5) OR: 1.140 (1.087 -1.197)			0.81 (0.78 - 0.84)	
Lairson *et al.* (2015)* (304)	USA	SEER - Medicare	17 US Registries 01/2006 - 12/2009 Age 65 - 94 Stage IIIB - IV n (PSM) = 4,884	Targeted Therapy & Platinum Based Therapy n (PSM) = 1,628 (33.3)	Platinum Based Chemotherapy/ No Chemotherapy	NR	Poverty Level	1st (Low) 290/863 (33.6)	2nd 400/1,256 (31.8)	3rd 425/1,242 (34.2)	4th (High) 513/1,523 (33.7)	1.00 (0.83 - 1.19)	6

Table 2.1 Continued

Sampling Frame						SES		Utilisation by SES Grouping (Number, %)					
Study	**Country**	**Data Source**	**Study Population[a]**	**Predictive Biomarker Test/Biological and Precision Medicine Overall Utilisation (Number, %)**	**Comparator**	**Unit**	**Measure**	**Lowest SES Group**			**Highest SES Group**	**Utilisation in Low SES (OR & 95% CI)[b]**	**QA**
Hepatobiliary Cancer: Novel Anti-Cancer Therapies													
Sahara *et al.* (2019)* (305)	USA	NCDB	1,500 Hospitals that Represent 70% of New Oncology Cases 01/2004 - 12/2015 Stage: I - IV n = 249,913 Hepatobiliary	Immunotherapy n = 585 (0.2)	No Immunotherapy	NR	Median Income	< $30,000 (Low) 79/40,313 (0.2) P Value: **0.006** Multivariate OR: Ref	$ 30,000 - $35,999 86/45,053 (0.2) Multivariate OR: 1.01 (0.74 - 1.38)	$36,000 - $45,999 157/67,104 (0.2) Multivariate OR: 1.23 (0.93 - 1.62)	≥ $46,000 (High) 242/87,867 (0.3) Multivariate OR: 1.43 (1.11 - 1.87)	0.71 (0.54 - 0.92)	8
Sanoff *et al.* (2016)*(306)	USA	SEER - Medicare	28% of US Population 2008 - 2011 Stage: Advanced Hepatocellular Cancer n =1,532	Sorafenib n = 422 (27.5)	No Treatment	Census Tract	% Below the Poverty Line	Q1 (Low) 104/335 (31.0)	Q2 99/368 (26.9)	Q3 109/414(26.3)	Q4 (High) 108/397(27.2)	1.20 (0.86 - 1.68)	6
Parsons *et al.* (2017) (307)	USA	SEER	NCI POC Study of SEER Patients 2007 & 2012 BCLC Stage C Hepatocellular Cancer n = 550	Sorafenib n = 186 (33.8)	NR	Census Tract	Median Income Per Year for the Individual's Census Tract	≤ $50,000 (Low) 62/264 (23.5) P Value: 0.06 OR: Ref	> $50,000 (High) 96/286 (33.6) OR: 1.50 (0.89 - 2.55) P Value: > 0.05			0.61 (0.41 - 0.90)	6.5

Table 2.1 Continued

Sampling Frame						SES		Utilisation by SES Grouping (Number, %)					
Study	**Country**	**Data Source**	**Study Population[a]**	**Predictive Biomarker Test/Biological and Precision Medicine Overall Utilisation (Number, %)**	**Comparator**	**Unit**	**Measure**	**Lowest SES Group**			**Highest SES Group**	**Utilisation in Low SES (OR & 95% CI)[b]**	**QA**
Sarpel *et al.* (2018)* (308)	USA	Electronic Medical Records	Urban, Tertiary Academic Healthcare and Leading Referral Center for Liver Diseases 2007 - 2013 Stage: Early which Progressed to Advanced (BCLC Stage C) Hepatocellular Cancer n = 959	Sorafenib n = 352 (36.7)	NR	Zip Code	Education (% with a Bachelors Degree)	< 24.3% (Low) 115/326 (35.3)	24.3% - 29.75% 82/261 (31.4)	29.75% - 42.3% 82/192 (42.7)	≥ 42.3% (High) 73/180 (40.6)	0.80 (0.54 - 1.18)	8.5
						Zip Code	Estimated Annual Median Income	< $37,309 (Low) 68/232 (29.3) OR (Higher vs Lower): 2.05 (1.19 - 3.54) P Value: **< 0.01**	$37,309 - $55,965 116/368 (31.5)	$55,965 - $ 82,814 84/215 (39.1)	≥ $82,814 (High) 84/144 (58.3)	0.30 (0.19 - 0.47)	
Melanoma: Novel Anti-Cancer Therapies													
Enewold *et al.* (2017)*[d] (323)	USA	SEER Hospital Medical Records, Contact with Treating Physicians and Others Involved Individual's Care	NCI POC Study of SEER Patients 2011 Age ≥ 20 Metastatic Melanoma n = 520[c]	Ipilimumab n = 109 (21.0)	NR	Census Tract	Median Household Income	$ 15,769 - $41,350 (Low) 22/124 (17.5)[c] P Value: 0.22	$41,351 - $55,155 21/128 (16.2)[c]	$55,156 - $73,178 30/130 (23.4)[c]	$73,179 - $163,393 (High) 33/133 (25.1)[c]	0.65 (0.34 - 1.25)	5
						Census Tract	% of Individuals Aged 25+ with at Least a High School Education	28.30% - 77.28% (Low) 24/126 (19.1)[c] P Value: 0.41	77.39% - 86.14% 22/126 (17.5)[c]	86.15% - 91.26% 27/133 (20.5)[c]	91.27% - 100.00% (High) 33/130 (25.2)[c]	0.69 (0.36 - 1.31)	

Table 2.1 Continued

Sampling Frame						SES		Utilisation by SES Grouping (Number, %)					
Study	**Country**	**Data Source**	**Study Population[a]**	**Predictive Biomarker Test/Biological and Precision Medicine Overall Utilisation (Number, %)**	**Comparator**	**Unit**	**Measure**	**Lowest SES Group**			**Highest SES Group**	**Utilisation in Low SES (OR & 95% CI)[b]**	**QA**
Al-Qurayshi *et al.* (2018)* (309)	USA	NCDB	Captures 70% of Newly Diagnosed Malignancies 2004 - 2012 Age ≥ 18 Stage III Cutaneous Melanoma n = 6,165[c]	Immunotherapy & Surgery n = 1,854 (30.1)	Surgery Only	Zip Codes	Community Household Income	Q1 (Low)	Q2	Q3	Q4 (High)	0.83 (0.69-1.01)	6.5
								202/746 (27.1)[c]	386/1,307 (29.5)[c]	523/1,702 (30.7)[c]	743/2,411 (30.8)[c]		
								P Value: 0.22					
								OR: 1.27 (0.99 - 1.62)	OR: 1.18 (0.98 - 1.43)	OR: 1.13 (0.96 - 1.33)	OR: Ref		
								P Value: **0.06**	P Value: 0.08	P Value: 0.13			
						Zip Codes	High School Graduation Rate	Q1 (Low)	Q2	Q3	Q4 (High)	0.65 (0.53 - 0.79)	
								178/758 (23.5)[c]	404/1,412 (28.6)[c]	595/1,893 (31.4)[c]	677/2,108 (32.1)[c]		
								P Value: **< 0.001**					
								OR: 0.59 (0.45 - 0.76)	OR: 0.80 (0.66 - 0.97)	OR: 0.94 (0.80 - 1.10)	OR: Ref		
								P Value: **<0.001**	P Value: **0.02**	P Value: 0.43			
Haque *et al.* (2019)*(310)	USA	NCDB	Captures 70% of Newly Diagnosed Malignancies 2004 - 2014 Age ≥ 18 Metastatic Melanoma n = 15,941	Immunotherapy n = 2,448 (15.4)	No Immunotherapy	Zip Code	Median Annual Income	< $63,000 (Low)	≥ $63,000 (High)			0.76 (0.70 - 0.84)	8.5
								1,482/10,418 (14.2)	929/5,202 (17.9)				
								P Value: **< 0.001**					
								OR: Ref	OR: 1.233 (1.118 - 1.360)				
								P Value: **< 0.001**					

Table 2.1 Continued

Sampling Frame						SES		Utilisation by SES Grouping (Number, %)					
Study	**Country**	**Data Source**	**Study Population[a]**	**Predictive Biomarker Test/Biological and Precision Medicine Overall Utilisation (Number, %)**	**Comparator**	**Unit**	**Measure**	**Lowest SES Group**			**Highest SES Group**	**Utilisation in Low SES (OR & 95% CI)[b]**	**QA**
Colorectal Cancer: Novel Anti-Cancer Therapies													
Fu *et al.* (2014)* (311)	USA	SEER - Medicare	17 US Registries (26% of US Population) 01/2005 - 12/2009 Age ≥ 65 Stage: IV & Early CRC with Progression or Recurrence n = 8,645[c]	Bevacizumab with Chemotherapy n = 4,502 (52.1)	No Bevacizumab	Census Tract	Median Household Income	Q1 (Low) 1,062/2,083 (51.0)[c] P Value: 0.66 Q1 OR: Ref P Value: **0.021**	Q2 1,112/2,101 (53.0)[c]	Q3 1,130/2,161 (52.3)[c]	Q4 (High) 1,175/2,265 (52.0)[c] Q4 OR: 1.20 (1.03 - 1.40)	0.96 (0.86 - 1.09)	8.5
						Census Tract	% of Adults with Less than a High School Education	Q1 (Low) 1,126/2,196 (51.3)[c] P Value: 0.20	Q2 1,126/2,187 (51.5)[c]	Q3 1,071/2,075 (51.6)[c]	Q4 (High) 1,162/2,153 (54.0)[c]	0.90 (0.80 - 1.01)	
Meyerhardt *et al.* (2012) (312)	USA	SEER - Medicare	16 US Registries (26% of US Population) 2002 - 2007 CRC Age ≥ 65 Stage IV n = 2,526	Bevacizumab with Combination Chemotherapy n = 903 (35.7)	Combination Chemotherapy Without Bevacizumab	Zip Code	Median Income	Q4 (Low) 214/631 (33.9)	Q3 229/631 (36.3)	Q2 229/631 (36.3)	Q1 (High) 231/632 (36.6)	0.89 (0.70 - 1.13)	5.5
Cen *et al.* (2012)* (313)	USA	SEER - Medicare	17 US Registries (26% of US Population) 2003 - 2005 Age ≥ 65 Stage I - IV CRC n = 46,692[c]	Chemotherapy Containing Bevacizumab Based Regimens n =1,306 (2.8)	No Chemotherapy/ Other Chemotherapy/ 5-FU Alone/ Oxaliplatin Based Chemotherapy Regimens	Census Tract	% of Residents Below the Poverty Line	4th (Low) 272/11,643 (2.3)[c]	3rd 337/11,669 (2.9)[c]	2nd 336/11,636 (2.9)[c]	1st (High) 360/11,662 (3.1)[c]	0.75 (0.64 - 0.88)	4

Table 2.1 Continued

Sampling Frame						SES		Utilisation by SES Grouping (Number, %)						
Study	**Country**	**Data Source**	**Study Population[a]**	**Predictive Biomarker Test/Biological and Precision Medicine Overall Utilisation (Number, %)**	**Comparator**	**Unit**	**Measure**	**Lowest SES Group**				**Highest SES Group**	**Utilisation in Low SES (OR & 95% CI)[b]**	**QA**
Parikh *et al.* (2016)* (314)	USA	SEER - Medicare	16 US Registries (28% of US Population) 01/2004 - 12/2009 Age ≥ 65 Stage IV CRC n = 4,418	First Line Chemotherapy with a Targeted Biologic n = 2,077 (47.0)	First Line Chemotherapy Treatment	Zip Code	% Below the Poverty Line	1st (Low) 457/989 (46.2)	2nd 504/1,070 (47.1)	3rd 549/1,139 (48.2)	4th (High) 567/1,220 (46.5)		0.99 (0.83 - 1.17)	7
Neugut *et al.* (2012) (315)	USA	SEER - Medicare	26% of US 01/2005 - 12/2005 Colon Cancer Age ≥ 65 Stage IV n = 859	Chemotherapy with Bevacizumab n = 310 (36.1)	Chemotherapy without Bevacizumab	Census Tract	SES Score based on education, poverty and income	1st (Low) OR: Ref	2nd OR: 1.98 (0.78 - 5.05)	3rd OR: 1.23 (0.50 - 3.01)	4th OR: 1.24 (0.48 - 3.26)	5th (High) OR: 2.30 (0.86 - 6.18)	[h]	6.5
Raab *et al.* (2019)* (316)	USA	SEER - Medicare	26% of US Population 01/2005 - 12/2013 Age ≥ 65 Stage IV Colon Cancer n = 3,785	Bevacizumab in Conjunction with Chemotherapy n = 2,352 (62.1)	Did Not Receive Bevacizumab in Conjunction with Chemotherapy	Census Tract	SES Rank (Education Level, Poverty Level & Income Combined Score)	Rank 0 (Low) 409/632 (64.7) P Value: > 0.05 OR: Ref	Rank 1 484/747 (64.8) OR:1.00 (0.79 - 1.28)	Rank 2 366/619 (59.1) OR: 0.87 (0.67 - 1.13)	Rank 3 608/994 (61.2) OR: 0.97 (0.76 - 1.23)	Rank 4 (High) 485/792 (61.2) OR: 1.13 (0.86 - 1.47)	1.16 (0.93 - 1.45)	9

Table 2.1 Continued

Sampling Frame						SES		Utilisation by SES Grouping (Number, %)					
Study	**Country**	**Data Source**	**Study Population[a]**	**Predictive Biomarker Test/Biological and Precision Medicine Overall Utilisation (Number, %)**	**Comparator**	**Unit**	**Measure**	**Lowest SES Group**			**Highest SES Group**	**Utilisation in Low SES (OR & 95% CI)[b]**	**QA**
Shih *et al.* (2009)[f] (297)	USA	NCDB	Captures 75% of Newly Diagnosed Cancer Cases 1998 - 2004 Stage: IV MCRC n = 16,027	Immunotherapy n = 662 (4.1)	NR	Zip Code	% Without a High School Degree	29%+ (Low)	20% - 28.9%	14% - 19.9%	< 14% (High)	[h]	7
								OR: 0.60 (0.35 - 1.03)[g]	OR: 0.70 (0.45 - 1.07)[g]	OR: 0.79 (0.53 - 1.16)[g]	OR: Ref[g]		
						Zip Code	Median Household Income	< $30,000 (Low)	$30,000 - $34,999	$35,000 - $45,999	$46,000+ (High)		
								OR: Ref[g]	OR: 1.09 (0.66 - 1.81)[g]	OR: 0.92 (0.56 - 1.51)[g]	OR: 0.92 (0.56 - 1.52)[g]		
Taylor *et al.* (2019)* (317)	USA	NCDB	Captures 70% of Newly Diagnosed Malignancies from 1,500+ CoC Facilities 2004 - 2015 Anorectal Melanoma n = 1,305	Immunotherapy n = 221 (16.9)	No Immunotherapy	Zip Code	Median Household Income	< $38,000 (Low)	$38,000 - $47,999	$ 48,000 - $62,999	≥ $63,000 (High)	0.87 (0.57 - 1.31) [k]	5.5
								>25/194 ([j])	47/305 (15.4)	68/359 (18.9)	74/426 (17.4)		
								P Value: 0.68					
						Zip Code	% High School Failure Rate	≥ 21.0% (Low)	13.0% - 20.9%	7.0% - 12.9%	< 7.0% (High)	0.95 (0.62 - 1.46) [k]	
								>25/212([j])	51/307 (16.6)	79/414 (19.1)	61/353 (17.3)		
Renal Cell Carcinoma: Novel Anti-Cancer Therapies													
Saigal *et al.* (2010)* (318)	USA	SEER - Medicare	15 US Registries 1992 - 2002 Stage: Metastatic n = 3,730[c]	IL-2 n = 560 (15.0)	Radial Nephrectomy/ Both Treatments/ Neither Treatment	Census Tract	Median Income	< $35,000 (Low)	$35,000 - $45,000	$45,000 - $60,000	> $60,000 (High)	0.77 (0.53 - 1.10)	4.5
								68/970 (7.0)[c]	70/1,007 (7.0)[c]	87/970 (9.0)[c]	70/783 (9.0)[c]		
								P Value: 0.3373					
						Census Tract	% of Non-High School Graduates	> 35% (Low)	20% - 35%	10% - 20%	< 10% (High)	0.72 (0.42 - 1.19)	
								22/373 (6.0)[c]	65/933 (7.0)[c]	141/1,567 (9.0)[c]	72/895 (8.0)[c]		
								P Value: **0.0394**					

Table 2.1 Continued

Sampling Frame						SES		Utilisation by SES Grouping (Number, %)				
Study	**Country**	**Data Source**	**Study Population[a]**	**Predictive Biomarker Test/Biological and Precision Medicine Overall Utilisation (Number, %)**	**Comparator**	**Unit**	**Measure**	**Lowest SES Group**		**Highest SES Group**	**Utilisation in Low SES (OR & 95% CI)[b]**	**QA**
Head & Neck Cancer: Novel Anti-Cancer Therapies												
Amini *et al.* (2018)* (319)	USA	SEER - Medicare	28% of US Population 2006 – 2011 Age ≥ 65 Oropharyngeal Cancer n = 409 28% of US	Cetuximab n = 173 (42.3)	Cisplatin/ Carboplatin	Census Tract	Median Income	Lowest (Low) 128/307 (41.7) OR (Cetuximab over Cisplatin) 1.86 (0.93 - 3.73) P Value: 0.079	Other (High) 45/102 (44.1) OR: Ref		0.91 (0.56 - 1.46	8
Xiang *et al.* (2018)* (320)	USA	SEER - Medicare	17 Registries (30% of US Population) 2004 - 2013 Age > 65 Stage III - IVB Oropharynx, Larynx & Hypopharynx Cancers n = 1,395 n (PSM) = 828	Cetuximab & Radiotherapy n = 609 (43.7) n (PSM) = 414 (50.0)	Cisplatin & Radiotherapy	Census Tract	Poverty Level	> 20% (Low) 128/293 (43.7) P Value: 0.22 PSM: 88/184 (47.8) P Value: 0.61	10% - 20% 154/384 (40.1) PSM: 111/228 (48.7)	< 10% (High) 327/718 (45.5) PSM: 215/416 (51.7)	0.93 (0.70 - 1.23) 0.86 (0.60 - 1.23)	6

Table 2.1 Continued

Sampling Frame						SES		Utilisation by SES Grouping (Number, %)						
Study	**Country**	**Data Source**	**Study Population[a]**	**Predictive Biomarker Test/Biological and Precision Medicine Overall Utilisation (Number, %)**	**Comparator**	**Unit**	**Measure**	**Lowest SES Group**				**Highest SES Group**	**Utilisation in Low SES (OR & 95% CI)[b]**	**QA**
Mixed Cancers: Novel Anti-Cancer Therapies														
Hershman *et al.* (2013)* (321)	USA	SEER - Medicare	Registry Coverage: NR 01/2004 - 12/2007 Age > 65 Stage IV or Recurrent Breast, Colon & NSCLC Cancers n = 16,085	Bevacizumab n = 3,039 (18.9)	No Bevacizumab	Census Tract	SES Score Based on Education, Poverty and Income	1st (Low)	2nd	3rd	4th	5th (High)	0.85 (0.74 - 0.98)	9
								327/1,870 (17.5)	514/2,901 (17.7)	689/3,590 (19.2)	697/3,642 (19.1)	810/4,065 (19.9)		
								P Value: 0.08						
								OR: Ref	OR: 0.97 (0.82 - 1.14) P Value: 0.75	OR: 1.04 (0.88 - 1.22) P Value: 0.65	OR: 1.06 (0.90 - 1.25) P Value: 0.53	OR: 1.09 (0.92 - 1.28) P Value: 0.33		
Mohile *et al.* (2013)* (322)	USA	Cancer and Aging Research Group's 'Determining the Utility of an Assessment Tool for Older Adults with Cancer' Multi-institutional Trial	7 Institutions 2006 – 2009 Age ≥ 65 CRC & NSCLC Scheduled to Receive a New Chemotherapy Regimen n = 207	Bevacizumab Plus Chemotherapy n = 27 (13.0)	Chemotherapy Alone	NR	Education	Less than 9th Grade (Low)	9th Grade/High School	Some College	Associate Degree + (High)		2.13 (0.27 - 13.4)	6
								3/12 (25.0)	10/80 (12.5)	9/78 (11.5)	5/37 (13.5)			
								P Value: 0.6984						

[a] Refers to the total number of patients in the cohorts of interest.
[b] Author generated.
[c] Numbers generated from percentage reported data which in some instances may lead to under or over estimations of true patient numbers in each SES measure sub-grouping as well as for the overall study numbers.
[d] Study reported twice in table due to reporting of both predictive biomarker test and biological and precision therapy data of interest.
[e] Data extracted for white women only as data for some black women was suppressed where number counts were ≤ 11.
[f] Study reported twice in table due to separation of results for two cancers of interest.
[g] OR data are for the ≥ 65 cohort. Data was unavailable to report the ≥ 65 total cohort study population numbers.
[h] Unable to calculate OR from raw data.
[I] TKI data prioritized for calculations and meta-analyses (larger population sample with oncogenic driver for therapy in question). This was to avoid potentially counting patients twice, should they have also received bevacizumab.

[j] Unable to calculate patient numbers receiving targeted treatments due to raw data suppression on account of small numbers. Authors have used a minimum value of 25 to represent >25 for total study population calculation purposes.
[k] OR calculated from comparison of Q2 to Q4.
* Eligible for inclusion in the meta-analysis.
All individual study OR reported are for predictive biomarker test and biological and precision therapy utilisation unless otherwise stated.
Quality appraisal scores could range from 0 (lowest) to 10 (highest).
Italicised & Bolded Studies = Paper selected for reporting when multiple publications reporting identical or heavily overlapping study populations presented.
P Values = Significant at $P < 0.05$
Abbreviations: BCLC: Barcelona Clinic Liver Cancer Stage; CALR: Cancer Activity Level Reporting Database; CCR: California Cancer Registry; CI: 95% Confidence Interval; CIHI-DAD: Canadian Institute of Health Information Discharge Abstract Database; CoC: Commission on Cancer; CRC: Colorectal Cancer; CRN: Cancer Research Network; ER; Estrogen Receptor; HER2: Human Epidermal Growth Factor Receptor 2; HMO: Health Maintenance Organisation; ICES: Institute for Clinical Evaluative Sciences; IL-2: Interleukin-2; IRR: Incidence Rate Ratios; NACRS: National Ambulatory Care Reporting System; NCCN: National Comprehensive Cancer Network; NCDB: National Cancer Database; NCI: National Cancer Institute; NDFP: New Drugs Funding Program; NR: Not Reported; NSCLC: Non-small Cell Lung Cancer; NYSCR: New York State Cancer Registry; OCISS: Ohio Cancer Incidence Surveillance System; OCR: Ontario Cancer Registry; ODB: Ontario Drugs Benefit; OHIP: Ontario Health Insurance Plan; OR: Odds Ratio; POC: Patterns of Care; PR: Progesterone Receptor; PSM: Propensity Score Matching; QA: Quality Appraisal; RPDB: Registered Persons Database; RR: Risk Ratio; SEER: Surveillance, Epidemiology and End Results Program; SES: Socio-economic Status; TKIs: Tyrosine Kinase Inhibitors.

2.4.3 Quality Appraisal

The 48 studies scored in the range 4-10, out of a possible 10 (mean = 6.9, median = 6.5) (Table 2.2). Papers scored well regarding data source(s), study populations and reporting of SES definition(s). For example, 40 out of 48 papers scored full marks for reporting complete SES data, including addressing any missing information (Question 1). A further 40 out of 48 papers also scored full marks for clear documentation of the SES unit measure (Question 2). Discussion of results with reference to the role of SES, statistical analysis with summary measures (e.g. ORs), and explanations for confounder selection were often reported poorly. For example, only 18 out of 48 studies discussed SES findings in the discussion (Question 9) and only 22 papers out of 48 reported adjusted analyses with confounder(s) listed (Question 8).

Table 2.2 Study quality appraisal results for the 48 included studies

Cancer	Paper	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10	Total
Predictive Biomarker Testing												
Testing	Pensa *et al*. (2009) (278)	0.5	1	0.5	1	1	0	1	0	0	0.5	5.5
	Lund *et al*. (2010) (279)	1	0.5	1	1	1	1	0	0	0	1	6.5
	Ferrusi *et al*. (2013) (280)	1	1	1	1	1	0.5	1	1	1	0	8.5
	de Camargo Cancela *et al*. (2015) (281)	1	0.5	1	1	0	0	1	1	1	0.5	7
	Greenbaum *et al*. (2017) (282)	1	1	1	0.5	1	0.5	1	1	1	0.5	8.5
	Rico *et al*. (2016) (283)	1	1	1	1	1	1	1	1	1	0.5	9.5
	Webster *et al*. (2013) (284)	1	1	1	1	1	1	1	1	1	1	10
	Enewold *et al*. (2017)[a] (323)	1	0.5	1	0	0	1	0	0	0	1	4.5
	Enewold & Thomas (2016)[a] (324)	1	0.5	1	0	0	1	0	0	1	0.5	5
	Palazzo *et al*. (2019[a] (325)	1	1	1	1	1	1	1	1	1	1	10
	Presley *et al*. (2018) (285)	1	1	1	0.5	1	1	0	0	0	0.5	6
Novel Anti-Cancer Therapies												
Breast	Du *et al*. (2011) (286)	1	0.5	1	1	0	0.5	0	0	0	1	5
	Vaz-Luis *et al*. (2015) (287)	1	1	1	0.5	1	0	0	0	0	0.5	5
	Reeder-Hayes *et al*. (2016) (288)	1	1	1	1	1	0.5	1	1	1	1	9.5
	Vaz-Luis *et al*. (2016) (289)	1	1	1	1	0	1	1	1	0	0.5	7.5
	Tsai *et al*. (2017) (290)	1	1	0.5	0	1	0	1	0.5	0	1	6
	Freedman *et al*. (2013) (291)	1	0.5	0.5	0	1	1	1	1	1	1	8
	Haas *et al*. (2011) (292)	1	1	1	0	1	0.5	1	1	1	1	8.5
	Goldhar *et al*. (2016) (293)	1	1	0.5	0.5	1	1	0	0	0	1	6
	Kumachev *et al*. (2016) (294)	0.5	0.5	1	0.5	0	1	0	0	0	1	4.5
	Thavendirenathan *et al*. (2016)[b] (295) & (2018) (327)	1	1	1	1	1	0	0	0	0	0	5
	Li *et al*. (2018) (296)	1	1	0.5	0.5	1	0.5	0	0	1	0.5	6
	Shih *et al*. (2009)[a] (297)	0.5	1	1	0.5	1	0	1	1	0	1	7
Lung	Zhu *et al*. (2012) (298)	1	1	1	1	1	1	0	1	0	1	8
	Langer *et al*. (2014) (299)	1	1	1	1	1	0.5	0	0	0	0.5	6
	Ritzwoller *et al*. (2014)[b] (300), Delate *et al*. (2014) (328) & Carroll *et al*. (2015) (329)	0.5	1	1	1	1	1	0	0	0	1	6.5
	Menter *et al*. (2016) (301)	0.5	0.5	0.5	1	1	0	0	1	0	1	5.5
	Maguire *et al*. (2019a) (326) & (2019b)[b] (302)	1	1	1	1	1	0	1	1	1	0.5	8.5
	Palazzo *et al*. (2019)[a] (325)	1	0.5	1	0.5	1	1	1	1	1	1	9
	Enewold & Thomas (2016)[a] (324)	0.5	0.5	1	0.5	1	1	0	0	1	0.5	6

Table 2.2 Continued

Cancer	Paper	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10	Total
Lung	Verma *et al*. (2019) (303)	1	1	0.5	1	1	1	1	0.5	0	1	8
	Lairson *et al*. (2015) (304)	1	1	0.5	0.5	1	0	0	1	0	1	6
Hepatobiliary	Sahara *et al*. (2019) (305)	1	1	0.5	0.5	1	1	1	0.5	1	0.5	8
	Sanoff *et al*. (2016) (306)	1	1	1	1	1	0	0	0	0	1	6
	Parsons *et al*. (2016) (307)	1	1	1	0	0	1	1	1	0	0.5	6.5
	Sarpel *et al*. (2018) (308)	0.5	1	1	0.5	1	1	1	0.5	1	1	8.5
Melanoma	Enewold *et al*. (2017)[a] (323)	1	1	1	0.5	0	1	0	0	0	0.5	5
	Al-Qurayshi *et al*. (2018) (309)	0.5	1	0.5	0.5	0	1	1	1	1	0	6.5
	Haque *et al*. (2019) (310)	1	1	1	1	1	1	1	0.5	0	1	8.5
Colorectal	Fu *et al*. (2014) (311)	0.5	1	1	1	0	1	1	1	1	1	8.5
	Cen *et al*. (2012) (313)	1	1	1	1	0	0	0	0	0	0	4
	Meyerhardt *et al*. (2012) (312)	1	1	1	0.5	1	0	0	0	0	1	5.5
	Parikh *et al*. (2016) (314)	1	1	1	0.5	1	0	1	0.5	0	1	7
	Neugut *et al*. (2012) (315)	1	1	1	0.5	0	0	1	1	0	1	6.5
Colorectal	Raab *et al*. (2019) (316)	1	0.5	1	1	1	1	1	0.5	1	1	9
	Shih *et al*. (2009)[a] (297)	0.5	1	1	0.5	1	0	1	1	0	1	7
	Taylor *et al*. (2019) (317)	1	0.5	1	0.5	1	1	0	0	0	0.5	5.5
Renal Cell	Saigal *et al*. (2010) (318)	1	0	1	1	0	1	0	0	0	0.5	4.5
Head & Neck	Amini *et al*. (2018) (319)	1	1	1	1	1	0.5	1	0.5	0	1	8
	Xiang *et al*. (2018) (320)	0.5	1	1	1	1	1	0	0.5	0	0	6
Mixed	Hershman *et al*. (2013) (321)	1	1	1	1	1	1	1	1	0	1	9
Cancers	Mohile *et al*. (2013) (322)	0.5	1	0.5	1	1	1	0	0	0	1	6

Q1: Data sources; Q2: Methods (study population and variables); Q3 & 4: Methods (operational definitions); Q5, 6, 7, & 8: Results and statistics; Q9 & 10: Discussion/conclusions.
[a]Paper reported twice in the table under different cancer sites or included data on predictive biomarkers as well as novel anti-cancer therapies.
[b]Paper selected for reporting when multiple publications reporting identical or heavily overlapping study populations presented.
Q: Question

2.4.4 Predictive Biomarker Testing

Eleven studies reported data of interest for five predictive biomarker tests (278–285,323–325). These biomarkers were: HER2 (n = 4) (278–281); KRAS (n = 3) (282–284); BRAF (n = 1) (323); and lung biomarkers (EGFR and/or ALK) (n = 3) (285,324,325). Ten of these studies were included in the meta-analysis (278–285,323,325). The eleventh paper (Enewold and Thomas, 2016) (324) was excluded due to overlap of SEER registry coverage by date with Palazzo *et al.* (2019) (325) for this biomarker type. The meta-analysis studies covered the following cancers: breast (4 studies) (278–281); colorectal (3 studies) (282–284); melanoma (1 study) (323); and NSCLC (2 studies) (285,325). The pooled OR for predictive biomarker test receipt for the lowest versus the highest category of SES was 0.86 (95% CI 0.71, 1.05; I^2 = 86%; 10 studies) (Figure 2.2) indicating that patients with a low SES were 14% less likely to receive biomarker testing than patients with higher level of SES, though this was not formally statistically significant (test for overall effect p = 0.14). This pattern was broadly consistent across cancer sub-groups (4 breast cancer studies, 3 colorectal cancer studies, 2 lung cancer studies and 1 melanoma study; test for subgroup difference p = 0.08) but the pooled OR was only statistically significant in colorectal cancer (OR 0.76, 95% CI 0.65, 0.88; 3 studies). Heterogeneity was considered high when all studies were combined (I^2 = 86%) and arguably, to be expected given that this meta-analysis brought together diverse studies (different cancers and biomarker tests, often with varied study methods). Whilst the high I^2 does increase the chance that these results have occurred by chance, the magnitude, direction of the effect, along with a significant p value would suggest that SES is still an important factor in predictive biomarker test utilisation.

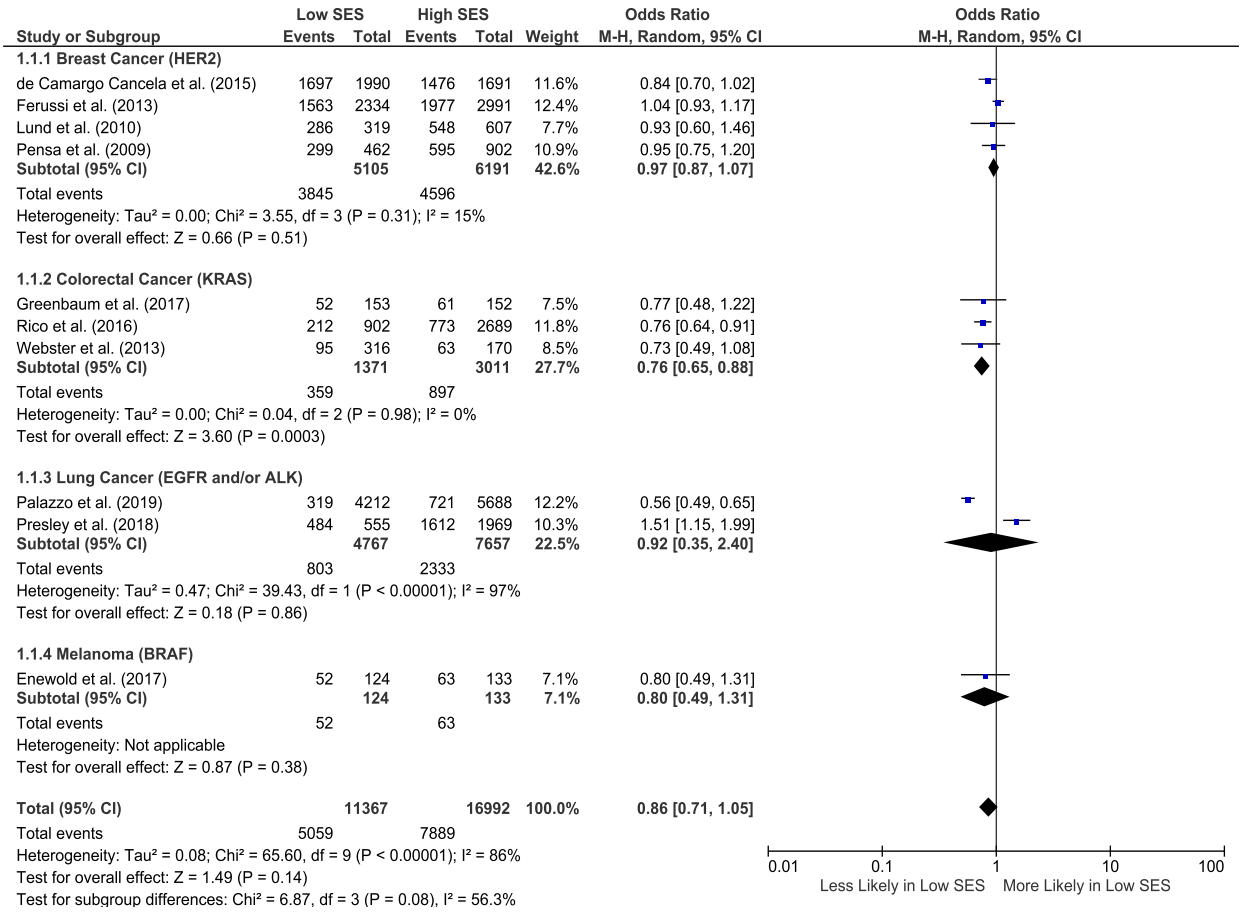


Figure 2.2 Forest plot showing predictive biomarker test utilisation odds (sub-grouped by cancer type) for low compared to high SES. Abbreviations: ALK: Anaplastic lymphoma kinase; BRAF: V-raf murine sarcoma viral oncogene homolog B1; CI: Confidence interval; EGFR: Epidermal growth factor receptor: HER2: Human epidermal growth factor receptor 2; KRAS: Kristen rat sarcoma virus; M-H: Mantel-Haenszel; OR: Odds ratio; SES: Socio-economic status.

2.4.5 Novel Anti-cancer Therapies: Primary Analysis

Associations of SES with novel anti-cancer therapy receipt was reported in 40 studies (286–325); thirty of these studies were included in the meta-analysis (286,288,290–292,294–296,299,301–306,308–311,313,314,316–323,325). The overall pooled OR for receipt of novel anti-cancer therapies for patients from low SES compared to those of high SES was 0.83 (95% CI 0.75, 0.91; I^2 = 85%; 30 studies) (Figure 2.3). The results from this unadjusted analysis indicate that patients with a low SES were 17% less likely to utilise a novel anti-cancer therapy than those patients with the highest level of SES and that this was statistically significant (test for overall effect p <0.05). Again, heterogeneity was noted to be high (I^2 = 85%), though the p value, magnitude, and direction of the effect of SES would imply that this was not just due to chance. Sub-group analyses I^2 tests were low for biologic therapies (I^2 = 60%) and immunotherapy (I^2 = 13%) too suggesting that studies considering targeted therapies (I^2 = 92%) may be the source of the variability.

Sub-group analysis suggested stronger associations with both immunotherapy utilisation (OR 0.82, 95% CI 0.78, 0.86; 7 studies) and targeted therapy (OR 0.80, 95% CI 0.62, 1.02) than for biological therapies (OR 0.91, 95% CI 0.81, 1.01). However, the test for sub-group differences was not significant (Figure 2.3).

Sensitivity analyses which substituted included studies for excluded studies with overlapping sampling frames confirmed the robustness of results (OR 0.80, 95% CI 0.72, 0.88; I^2 = 86%; 30 studies) (Appendix 2.10). Similar results were also observed in sensitivity analyses when only studies from the USA were considered (OR 0.82, 95% CI 0.74, 0.91, I^2 = 85%, 27 studies) (Appendix 2.11).

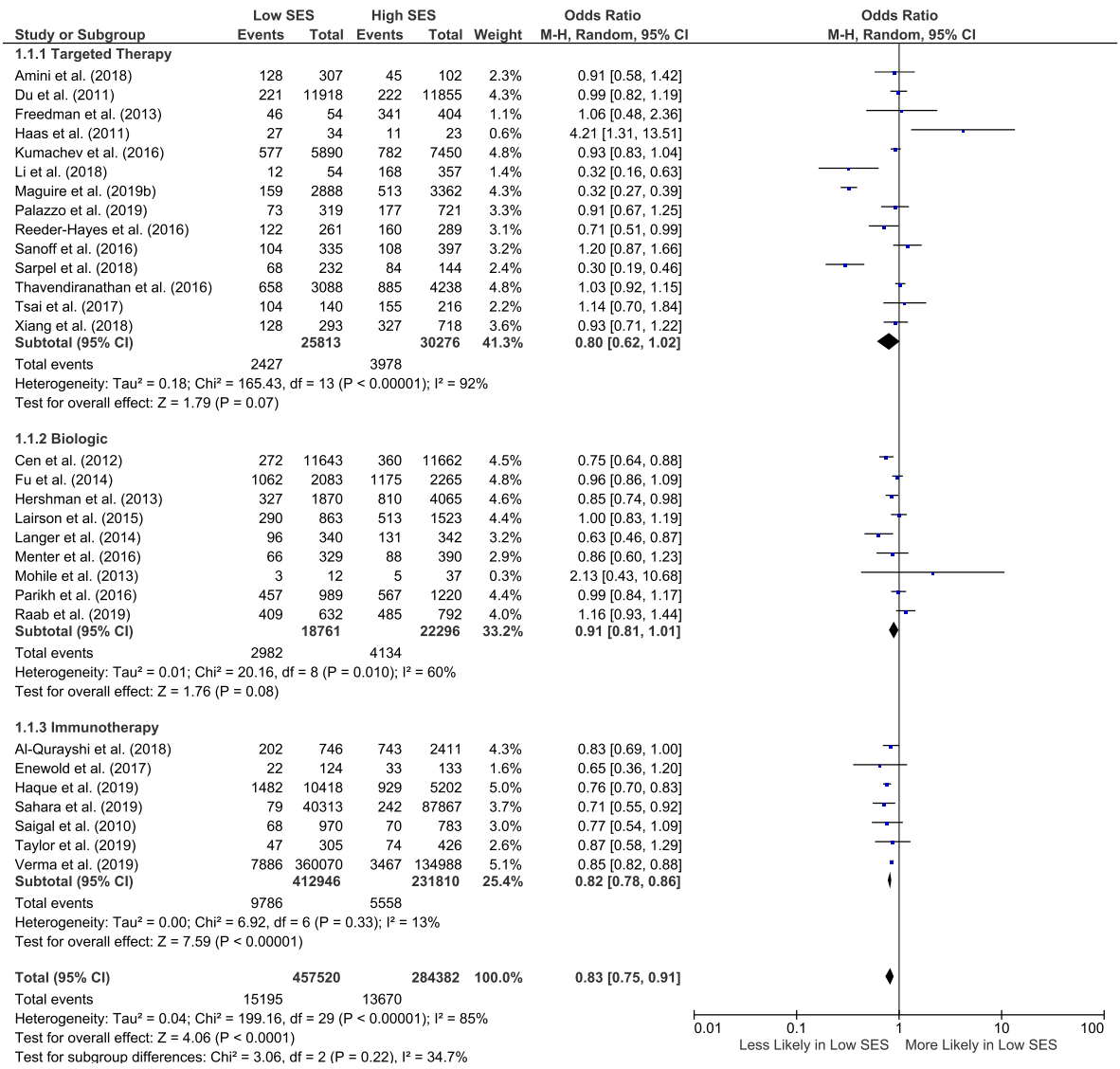


Figure 2.3 Forest plot showing novel anti-cancer therapy utilisation odds for all cancers (sub-grouped by drug class) for low compared to high SES. Abbreviations: CI: Confidence interval; M-H: Mantel-Haenszel; OR: Odds ratio; SES: Socio-economic status.

2.4.6 Novel Anti-cancer Therapies: Sub-group Analyses

For breast cancer, eleven studies reported the association of SES with the receipt of the HER2 targeted MAB, trastuzumab (286–296) and one with immunotherapy (297). Eight studies (all reporting trastuzumab utilisation) were eligible for meta-analysis (286,288,290–292,294–296). The pooled OR for receipt of trastuzumab in those with low, compared to high, SES was 0.93 (95% CI 0.78, 1.10; I^2 = 68%) (Figure 2.4).

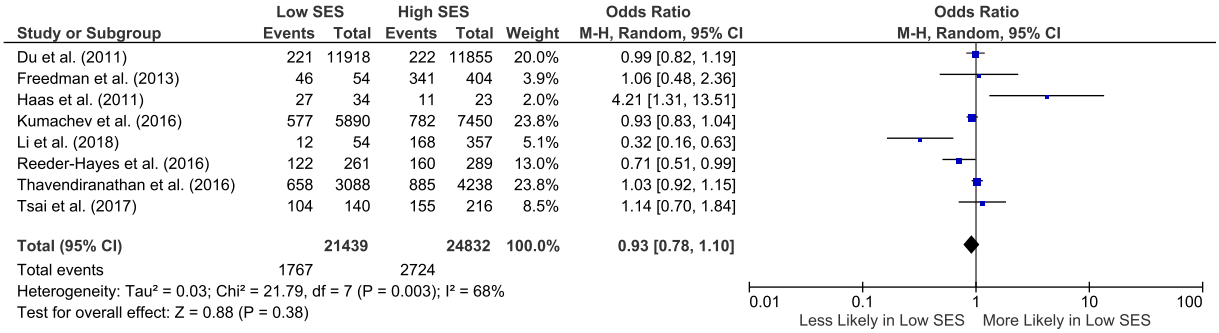


Figure 2.4 Forest plot showing novel anti-cancer therapy utilisation odds in breast cancer for low compared to high SES. Abbreviations: CI: Confidence interval; M-H: Mantel-Haenszel; OR: Odds ratio; SES: Socio-economic status.

Nine lung cancer studies evaluated SES with novel anti-cancer therapy receipt (298–304,324,325). Four of these reported: bevacizumab (298–301); two TKIs (324,325); one both bevacizumab and TKIs (302); one immunotherapy (303); and one biological therapies (mostly bevacizumab) (304). Six studies were eligible for meta-analysis (299,301–304,325), and the pooled OR for receipt of a biological or novel anti-cancer therapy in those of low, compared to high SES was 0.71 (95% CI 0.51, 1.00; I^2 = 95%) (Figure 2.5).

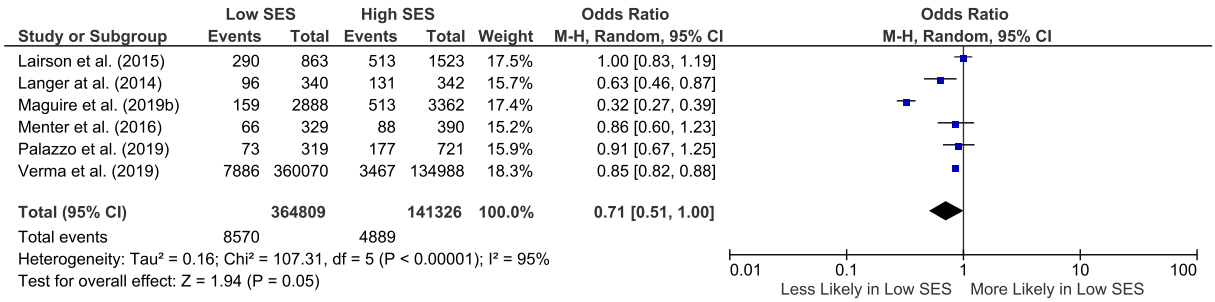


Figure 2.5 Forest plot showing novel anti-cancer therapy utilisation odds in lung cancer for low compared to high SES. Abbreviations: CI: Confidence interval; M-H: Mantel-Haenszel; OR: Odds ratio; SES: Socio-economic status.

Twenty studies reported data of interest for six other cancers: hepatobiliary (4 studies) (305–308); melanoma (3 studies) (309,310,323); colorectal (8 studies) (297,311–317); renal cell carcinoma (1 study) (318); and head and neck cancer (2 studies) (319,320). A further two studies reported data on more than one cancer (321,322). Studies referenced the following seven treatments: immunotherapy (297,305,309,310,317); bevacizumab (311–313,315,316,321,322); sorafenib (306–308); ipilimumab (323); targeted biologics (314); interleukin-2 (IL-2) (318); and cetuximab (319,320). Sixteen studies were combined into a meta-analysis (305,306,308–311,313,314,316–323), giving a pooled OR for receipt of novel anti-cancer therapies for low SES compared to high SES of 0.84 (95% CI 0.76, 0.94; I^2 = 73%) (Appendix 2.12). The test for sub-group differences between breast, lung, and all other cancers was not statistically significant suggesting that cancer type does not modify the effect of novel anti-cancer therapy by SES. However, as a relatively small number of lung and breast cancer studies contributed data to this analysis, this may mean that it is underpowered to detect subgroup differences (Appendix 2.13).

Finally, among all analyses, there was no clear observable evidence of publication bias (Appendix 2.14).

2.5 Discussion

2.5.1 Main Findings

This is the first systematic review and meta-analysis to examine whether there are inequalities in novel anti-cancer therapy utilisation and/or associated testing by SES. Overall, the findings show that there are statistically significant socio-economic inequalities in novel anti-cancer therapy utilisation; those with a low SES were 17% less likely to be treated with novel therapies compared to those with a high SES. An effect of similar magnitude was observed in test receipt but did not achieve statistical significance.

2.5.2 Interpretation of Findings

The finding that differences are present in novel anti-cancer treatments is consistent with previous systematic reviews documenting conventional treatment inequalities in cancer therapy (221,330). It was not possible to formally compare this meta-analysis to others previously published (of which there are very few reviews which report treatment utilisation ORs for a measure(s) of SES). However, it was noted, that where did this occur, similar effect magnitudes were observed with lung cancer chemotherapy utilisation (low versus high SEP; 0.82 (95% CI 0.72, 0.93; I^2 = 67%). Although the magnitude of the effect was stronger when surgery was considered (low versus high SEP; 0.68 (95% CI 0.63, 0.75; I^2 =53%) (221). Similar socio-economic inequalities have also been observed across the cancer care pathway from screening (331), to diagnosis (332), and timeliness of referral and treatment receipt (333), through to survival (143). Combined, this suggests that low SES remains a barrier to treatment access and cancer care, despite advances in treatment.

The strength of socio-economic inequalities varied with cancer type; the effect estimate for receipt of novel anti-cancer therapies was stronger for lung cancer (incidence of which is related to low SES) than other cancers. It is not clear why this is so, although the risk of some cancers (including lung) is associated with certain health behaviours (e.g. smoking) (334). It is possible that these health behaviours, alongside other factors (which, themselves may be a consequence of the health behaviours), such as multi-morbidity, could influence a healthcare professional's decision to offer or initiate cancer treatment (335). Additionally, patients may choose to refuse treatment and this has been shown previously in a stage IV NSCLC cohort to be influenced by socio-economic factors (336). Whilst such individual behavioural factors warrant further

investigation, they need contextualising within the wider determinants of health (i.e. the social, economic, cultural and clinical level factors) which are also associated with known treatment barriers (30). For example residing in areas of greater socio-economic deprivation has been linked to later stage cancer diagnosis and delayed cancer screening (122) - factors which could then result in further care barriers such as limited treatment options and/or treatment success (337).

Socio-economic inequalities in novel anti-cancer therapy utilisation in breast cancer were less pronounced than in lung, hepatobiliary, melanoma and renal cell cancers (Appendix 2.12), despite most of the published research focusing on breast cancer to date. A previous meta-analysis of 5 studies concluded equivocal associations between trastuzumab uptake and SES given a pooled estimate of economically advantaged patients in comparison to deprived patients of 1.03 (95% CI 0.86, 1.25; I^2 = 43%) (236). Whilst there is some overlap in the studies synthesised, this review included an additional 4 studies, more patients (n = 46,271 versus n = 4,294) and was more up to date (December 2019 versus July 2017). Direct comparisons of the two analyses are not fully feasible given that it is unknown which measure of SES the Martin *et al.* (2018) review used in situations when more than one measure was recorded in the study. However, when taken together, the differences listed here could explain why this review found a slightly stronger association by SES with regards to trastuzumab utilisation in breast cancer (236). Both review findings suggests that low SES may be less of a treatment barrier in breast cancer than in other cancers, at least as far as newer therapies are concerned. One possible explanation for this may be that breast cancer sub-type differentiation and the practice of hormone receptor status testing, and basing treatment on these results, is well established and routinely embedded in clinical practice (originating in the 1970s following the discovery of the ER) (338). Hence, these findings support the wider concept of the Inverse Equity Hypothesis (230); that is, whilst new interventions may temporarily widen inequalities by disproportionately favouring those with resources enabling priority access, over time this narrows as treatment access "trickles down" and becomes standard clinical practice (56,339).

In relation to predictive biomarker testing more generally, the observation that there is reduced utilisation with respect to lower SES builds on previously documented relationships between factors associated with SES and test receipt (e.g. negative association between smoking and EGFR and ALK abnormalities) (340). Previous work also highlights that test patterns vary temporally and spatially (341,342), as well as with respect to patient demographics (e.g. age) (343). This suggests that access to biomarker testing is complex. Nevertheless, the observation

that low SES may reduce access to testing has important implications. First, utilisation barriers occur at points other than just therapy receipt, a finding echoed by CRUK who highlighted that many colorectal and NSCLC patients potentially eligible for targeted treatments did not receive molecular biomarker testing (344). Second, if multiple barriers to novel therapy utilisation exist, then sophisticated solutions are likely required to prevent cancer inequalities widening further. This idea is discussed further in Chapter 5.

2.5.3 Strengths and Limitations

This is the first comprehensive meta-analysis on this important and growing area of practice and brings together data on over one million patients. Despite this, the study had several limitations.

It is acknowledged that there are challenges comparing studies reporting different measures of SES. There was no one consistent measure used which may be important given that different measures may be more (or less) appropriate for different populations, depending on their age, gender and context and the purpose for which SES is being assessed. An example of such a measure is household income and whether this accurately reflects SES for all household members, especially if they do not all have an equal share of the income (e.g. females) (345). Additionally, for the retired population, household income does not necessarily capture the life course exposures to deprivation and privilege which will also impact SES (346). Even when studies appeared to use the same measure (e.g. income) how the variable was categorised (e.g. what was considered "high" or the number of subgroups considered) differed. For most studies, there was considerable variation between what was considered "high" and "low" SES, meaning that true differences were unlikely to be attenuated by a lack of variability as cross study comparisons were difficult. However, as almost all studies used area-based socio-economic measures (termed "ecological" as the measure is at the level of a group), the ecological fallacy in inference is a risk (347). This means that it was not possible to make seemingly natural assumptions that associations seen at an area level SES measure also pertain to individuals with similar SES living within these groups (347). Measurement of SES in the studies contained in this analysis may therefore be subject to misclassification. To fully understand the impact of SES on the individual with regards to novel anti-cancer therapy utilisation, data collection ideally needs to be collected at this level, which in most instances is currently not feasible (healthcare data often lacks individual information for release due to privacy reasons). The limitations of SES measurement (including the ecological fallacy) are discussed throughout this thesis and a detailed critique is provided in Chapter 5.

Further limitations were also noted regarding results interpretation. First, determining ORs from raw data disregards adjustments for confounders; this along with variations in study sampling frames may, in part, explain the high heterogeneity observed. Although this may suggest that some results may have occurred by chance, the magnitude, direction of the effect, along with in some instances significant p values, would still suggest that SES is an important factor in both predictive biomarker test and novel anti-cancer therapy utilisation. Some studies (n = 24) did report adjusted OR data where important confounders of associations had been taken into consideration; had all studies done the same, this would have negated the need to have calculated unadjusted OR for inclusion in the meta-analysis. It does, however, mean that the possibility cannot be entirely excluded that any associations seen in the meta-analyses could be explained by uncontrolled confounding. Second, where conducted, the meta-analysis did not take into consideration study quality, and this may have had an impacted validity. Third, sub-group analyses may be misleading; multiple group comparisons can yield false positives (348) and inadequately powered analyses (sometimes considered to be less than 5 studies) (349) will be unable to detect effects of interest from random-effect models. Hence all sub-group analyses reported in this Chapter should be interpreted with caution. Fourth, whilst it is desirable to declare sensitivity analyses *a priori* to restrict the later integration of bias into the findings, this was not feasible in all instances. To minimise risk, clear rationales for *post hoc* analyses were provided though running analyses in this way does mean that it was not possible to remove entirely the possibility that findings could occur by chance.

Limitations were also noted in the review process. First, single reviewer title and abstract screening, while considered acceptable by the Cochrane Collaboration (255), may have erroneously excluded relevant studies. Ideally, two reviewers in duplicate would screen all records from the title and abstract stage. However, owing to larger search retrieval numbers (10,722 records) this was not considered feasible. Second, publications not written in English, those with only published conference abstracts and only querying study eligibility with authors at the data extraction stage, may have further contributed to the possibility that this review does not capture all evidence exploring SES in novel anti-cancer therapy use. Likewise, the absence of grey literature inclusion, in particular searching reports from cancer registries, also reduces the likelihood of this evidence base being both balanced (increased risk of publication bias as null findings are less likely to be published) and timely (time lag between submission to publication of evidence) (350,351). Moreover, papers in languages other than English were excluded due to a lack of resources for translation. It is therefore possible that some relevant studies were not included and potentially, these finding (if included) may have skewed the

results of the meta-analyses in either direction depending on their findings. However, steps were taken to reduce the risk of missing eligible papers as the single screener (RN) was conservative in their approach and erred on the side of caution, selecting for full text review any paper that could potentially be relevant (as evidenced by the fact that 551 citations proceeded to the full text review). Independent double screening of the full text was undertaken as a second check. Furthermore, at the time the search was conducted, there was not a requirement to report all databases, registries, and website grey literature searched as the change in PRISMA guidance now stipulate to be good reporting practice (243); thus, these sources were not specified on the PRISMA. Third, there is the potential for random error in healthcare data collection (e.g. incorrect reporting of drug utilisation in some patient instances) - though the large-scale nature of these finding should reduce this risk. Finally, there is always the possibility of undetected systematic error which could operate in either direction.

The review also highlights limitations in the evidence base. For example, whilst undesirable, low frequency of utilisation was reported in some studies; this may be reflective of the field given the rarity of some mutations these therapies target (for example, in NSCLC only around 1% of patients have been found to have a ROS1 gene coding rearrangement) (352). Additionally, the lack of published data on these newer tests and associated treatments may be indicative of pace of change within the precision medicine field. No analyses of time effects in precision medicine utilisation were undertaken for this review, and this is a further limitation of this work. However, it is possible to explore informally whether there are any suggestions of time effects. For example, though the review search window commenced 1998, the earliest studies of note were dated from 2009 (n=2). The search was run in 2019 and seven eligible papers were published that year and six in 2018. Increasing publication numbers in later years likely reflects increased uptake (and research interest) in novel anti-cancer therapies over time. It has been noted that patterns of novel anti-cancer therapy utilisation may differ in early versus later adopters (e.g. non-selective EGFR use in NSCLC) (353). For many of these treatments, because they are so new, it might be argued that all the reported data in this review relates to early adopters. Therefore, it would be informative and important to undertake a comprehensive investigation of time effects in any update to the review.

This review has also raised attention on the drawbacks of focusing on one type of healthcare system (non-publicly funded) - and whether the measure of SES reported in this analysis fully reflect the characteristics of such population in question. For example, income was the most common SES measure reported - yet this variable is known to vary across the life course. Hence,

SES measures such as median household income may be less meaningful in retired SEER populations (19). In such circumstances, eligibility for Medicare may be more important in addressing one of the most important barriers to care in the USA - that of having health insurance. Similarly, as employment is often tied to insurance coverage in non-publicly funded healthcare systems like the USA, this choice of socio-economic measure could be an additional factor related to utilisation outcomes in the under 65 age group other than income alone. The generalisability of conclusions drawn to patients outside the USA and age groups younger than 65 years must be questioned. Having said this, studies from other countries documented similar patterns in inequality (281,294,296). Moving forward, consideration of other registries known to be rich in novel anti-cancer therapy data would be valuable (354). The SEER registry also underrepresents minority populations. This limitation may be important given the links between ethnicity and genetics (128,355,356). For example, as EGFR mutated lung cancers are known to be associated, in part, with an Asian ethnicity (357), this demographic would be anticipated to be more likely to have a targetable EGFR mutation with increased sensitivity for EGFR inhibitor application. However, SEER registry data on Asian ethnicity is both rarely reported and biased (358). This means that SEER studies may not accurately reflect molecular targeted therapy application in this demographic and may underestimate the true the size effects of inequalities seen in some novel anti-cancer therapies applications.

2.5.4 Post Publication Update

Precision medicine practice and research is a fast-moving area. In the time that has elapsed since the search for the systematic review was completed (December 2019), multiple additional studies exploring associations between predictive biomarkers and novel anti-cancer therapies have been published. To keep abreast with these changes, an update was run through to April 2022. The findings are summarised below and provide a contemporary context for the interpretation of the thesis's empirical findings.

There has been additional work exploring the role of predictive biomarker testing within a NSCLC context; these studies have considered: EGFR (n = 2 studies) (353,359); PD-L1 biomarkers (n = 1 study) (360); and a mix of NSCLC biomarkers including EGFR and PD-L1 (n = 1 study) (361). Findings of these studies, with regards to SES inequalities in treatment utilisation, were mixed. Three studies (353,359,360) are consistent with the findings of the review that, biomarker uptake is lower in those considered to be of a lower SES. The fourth study (361) suggests that the role of SES depends on the specific biomarker and measure of SES used to quantify the association. With regards to NSCLC novel anti-cancer therapies, seven

additional studies (353,360,362–366) have explored the role of the following treatments: ICIs (n = 3 studies) (360,362,364); TKIs (n = 5 studies) (353,360,363,365,366); and bevacizumab (n = 1 study) (360). Four studies were conducted in the USA (353,362–364); one study in Canada (366); one study in China (365); and one study in Israel (360). Most of these studies reported the same trends in utilisation as observed in Chapter 2 and Chapter 4. However, two studies results did not show associations of reduced therapy utilisation with SES (353,360).

One study has explored the role of predictive biomarker testing within a breast cancer context (HER2 testing) but found that the odds of receiving testing were not associated with socio-economic factors of education and employment status (367). Four breast cancer treatment utilisation studies have been published - two of these concern data from the USA (368,369); one study data is from China (370); and the other from Singapore (367). Trastuzumab remained the drug of focus (367,368,370), though work on other anti-HER2 agents, including pertuzumab, is emerging (368,369). Generally, these studies reiterate the findings of Chapter 2 and Chapter 3; that is, reduced drug utilisation when SES is lower (although the magnitude of this effect is modest).

2.6 Conclusions

This Chapter determined that, for studies published up until December 2019, there are socio-economic inequalities in the utilisation of both predictive biomarker tests as well as novel anti-cancer therapies although only the latter reaches statistical significance. The degree of the associations for low SES with reduced novel therapy utilisation varied by cancer type and was found to be stronger in lung than breast cancers. The implications of these findings are discussed again, where appropriate, in the subsequent secondary linked data analysis studies (Chapters 3 & 4). Synthesised study data covered a range of solid tumour cancers and predictive biomarker tests, despite primarily focusing on the American non-publicly funded healthcare system. To resolve this limitation of the evidence base, a focus on population-based studies outside of a USA setting and where economic factors are less dominant is needed. One such database providing opportunity for such an analysis is England's SACT dataset - the analysis of this dataset in a breast cancer demographic, provides the basis for next Chapter.

Chapter 3. Socio-economic Inequalities in HER2+ Breast Cancer Trastuzumab Utilisation

3.1 Introduction

Chapter 1 introduced how it is widely accepted that a low SES is associated with reduced utilisation of conventional cancer treatments. Chapter 2 added to this knowledge by providing evidence for a growing body of evidence documenting similar reduced utilisation in low SES patients with newer, novel anti-cancer therapeutics (precision medicines, biologicals, and immunotherapies) and their associated predictive biomarker tests. However, to date this evidence is primarily focused on USA data (Appendix 2.1) (231). Only 5 published studies have explored associations of a low SES on novel anti-cancer therapy utilisation within publicly funded healthcare systems (272,293–295,327). Hence, the potential for SES to be a barrier to novel anti-cancer therapy utilisation in healthcare systems without the additional complexity of individual finance and health insurance (private or otherwise) remains underexplored. This Chapter sought to address this knowledge gap by investigating socio-economic inequalities within the context of a female, HER2+ breast cancer population eligible to utilise trastuzumab. It sought to compare whether patterns of utilisation of trastuzumab differed from those for conventional treatment in the same patients. This comparison could shed light on areas of SES inequalities.

Chapter Aim: Determine the association of SES (measured in terms of the IMD income domain of area of residence at diagnosis) on trastuzumab and conventional treatment (breast cancer directed surgery and chemotherapy) utilisation for a HER2+ breast cancer population in England.

Objectives:

1) Conduct a retrospective observational cohort study using English cancer registry data linked with the SACT dataset.
2) Undertake a primary analysis to determine likelihood of trastuzumab utilisation by SES for all patients.
3) Undertake a secondary analysis to determine likelihood of conventional treatment (breast cancer directed surgery and chemotherapy) utilisation by SES for all patients.

This Chapter provides the background to this study along with the steps taken to acquire, prepare and analyse the data. Results are then provided and discussed. The Chapter ends with concluding comments.

3.2 Background

Breast cancer was one of the first cancers over the past few decades to benefit from the increasing integration of novel anti-cancer therapies (371). One clear example is the discovery of the MAB, trastuzumab, used in patients with HER2+ breast cancer. Trastuzumab targets HER2/*neu* proto-oncogene amplification (372) and its introduction provided a treatment option beyond conventional choices (e.g. surgery and cytotoxic chemotherapy) (373) which improved the prognosis of what was traditionally considered an aggressive breast cancer sub-type (374). Subsequently, further novel anti-HER2 targeted agents (e.g. lapatinib, pertuzumab, and trastuzumab emtansine) have been developed (375) and integrated into treatment guidelines (376,377). At present, HER2+ breast cancer treatment is increasingly personalised and continually evolving. Despite numerous advances, trastuzumab still remains the standard of care in both early and metastatic HER2+ disease.

Trastuzumab utilisation data from England at a population level has seldom been reported, despite the therapy being the second most investigated novel anti-cancer therapy in Chapter 2 (Appendix 2.1) (231). Since the initial search for the systematic review was conducted, a further three observational studies have been published exploring trastuzumab utilisation - and these report similar socio-economic gradients in utilisation (367–369). There is some additional emerging evidence exploring socio-economic inequalities in novel anti-cancer treatment utilisation from China (a mixed healthcare system) (370) as well as Australia (publicly funded healthcare systems (378) but these studies use comparatively small cohorts with no denominator population which did not allow for assessment of the odds of utilisation. As a high-cost medicine, utilisation of trastuzumab in a publicly funded system is needed to explore if inequalities persist when private finance and insurance is not a factor in receipt. Given the previously documented low SES associations with conventional surgery and chemotherapy utilisation in breast cancer, it is important to understand if the novel anti-cancer therapy trastuzumab is also subject to similar socio-economic patterning. To investigate, a large, population-based observational cohort study in a publicly funded healthcare system (NHS in England) was undertaken. The aim of this study was to determine the association of SES with utilisation of: (i) trastuzumab; and (ii) comparative conventional treatments (breast cancer directed surgery and chemotherapy) in a HER2+ breast cancer population in England.

3.3 Materials & Methods

This population-based retrospective cohort study used English National Cancer Registry Data (NCRD) linked with the Systemic Anti-Cancer Therapy (SACT) dataset. Details on both these data sources are now described below.

3.3.1 NCRD

The National Cancer Registration and Analysis Service (NCRAS) collates, analyses and prepares for release the population-based cancer registry for England (379). The NCRD covers all people, living in England and diagnosed with malignant and pre-malignant neoplasms since 1971, detailing approximately 300,000 tumour records reported each year (379). NHS funded activity in the private sector is included, though as the majority of hospital activity is funded by the NHS (98-99%), information on private sector cancer registration is deemed incomplete (379). All registry data are legally curated under the provisions of Section 251 of the NHS Act 2006. This Act grants access to cancer patient data without consent across 162 healthcare providers (380). Historically, data was collated by eight regional English cancer registries, though this process has since been standardised with the formation of one national registry in 2012 (379). Registry data can be used to monitor cancer incidence, ensure quality assurance across the cancer care pathway and improve care (better diagnosis and treatment) (380).

The registry uses an event-based registration model which compiles data inputted from multiple hospital (e.g. MDT meetings, pathology reports, molecular test results, and treatment records) and other national data collection sources (e.g. patient waiting times, screening, and mortality records) (379). Data flow is now predominantly electronic in nature, with patient NHS numbers providing a unique identifier (ID) for data linkage (379). Data flow from NHS Trusts into the cancer registry uses a cancer management system (e.g. Somerset and Infoflex) (381). In 2013 a new, national standard for reporting cancer in England (the Cancer Outcomes and Service Dataset (COSD)) was introduced to replace the previous National Cancer Dataset (379,382). The COSD specifies the items required for submission to NCRD and as recording is in one record (e.g. information obtained from MDT software, patient administration systems (PAS), and pathology systems), it aimed to minimise the data collection burden on Trusts (382). At the time of this work, monthly COSD extracts were submitted to PHE (now NHS Digital)[4] for linkage by cancer registration officers with other data sources recorded on different systems (381). There are two COSD data submission per Trust (the COSDv9 dataset uploaded using the

[4]PHE was dissolved in 2021 and NHS Digital is now the registry data controller (390).

online secure portal and the other, a COSDv4 pathology submission which is sent via secure email). PHE (and now NHS Digital) therefore did not extract data themselves from Trust systems. Data linkage requires automated tools as well as manual extraction and review processes to ensure that records are linked correctly, and duplicates removed (379). The resulting full dataset (The English National Cancer Online Registration Environment (ENCORE)) provides the cancer registration record. From this, the Cancer Analysis System (CAS) clones this data to provide as follows: monthly snapshots for analysis, linkage with other NCRAS datasets, and data release. The process of this cancer registration data flow is depicted in Figure 3.1.

Released NCRD data are recorded in three tables (patient, tumour, and treatment) (379). The patient table comprises patient level information items (e.g. sex, ethnicity, age, number of comorbidities, and deprivation of area of residence), the tumour table describes primary tumour data summaries (e.g. stage, grade, and histology) and finally the treatment table details data available on tumour treatment (e.g. type, date and healthcare provider) (379). In the NCRD, deprivation is recorded as the income domain of IMD and grouped into quintiles. IMD represents a measure of relative deprivation for small areas (or neighbourhoods) in England (383).

English cancer registration is as follows: at the tumour level (which means that multiple tumours per individual may be registered); reports only primary cancers (i.e. no secondaries); follows international data recording conventions; and is subject to rigorous quality assurance checks. The main cancer registry tables can be linked with other mandated NCRAS datasets to expand the tumour level detail captured. The new SACT dataset provides one example of a linked dataset - and it was this dataset that was the focus of this work.

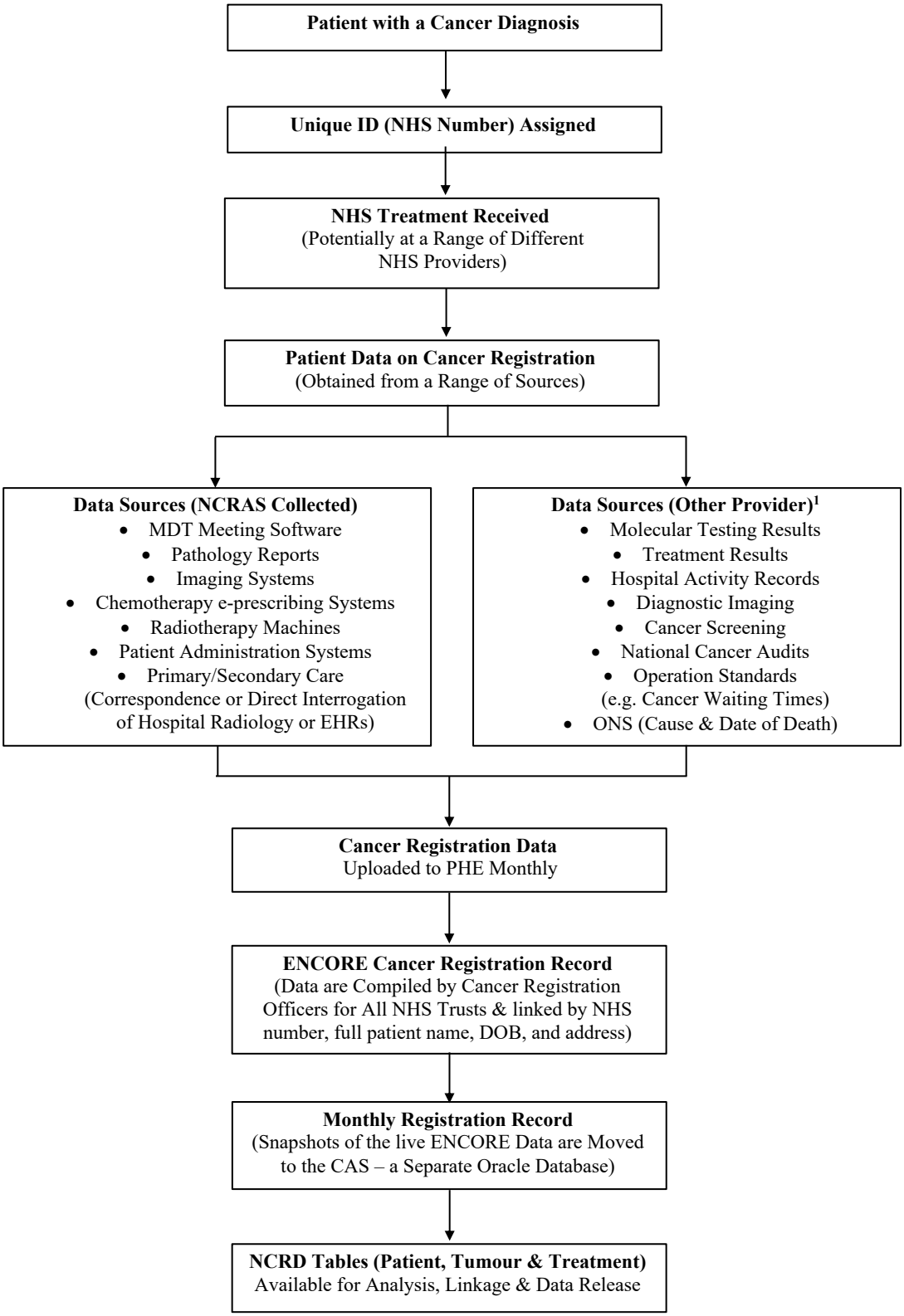


Figure 3.1 NCRD Data sources, data flow, and data linkage.
Please note, although all these data items may be collected, that does not mean that they are complete and/or made available to those requesting data for research purposes.
[1]Data on privately treated patients is not routinely submitted, however private hospitals may request patient consent to submit data to the cancer registry in some circumstances.
Abbreviations: CAS: Cancer Analysis System; DOB: Date of birth; EHRs: Electronic healthcare records; ENCORE: English National Cancer Online Registration Environment; ID: Identifier; NCRAS: National Cancer and Registration Analysis Service; NCRD: National Cancer Registry Database; NHS: National Health Service; MDT: Multidisciplinary team; ONS: Office for National Statistics.

3.3.2 SACT Database

The SACT database is a relatively new resource recording drug level information on routinely administered systemic anti-cancer therapy; it captures data from secondary and tertiary NHS providers when treatment is across day case, inpatient, outpatient, and community settings (354,384). The database provides details on adult and paediatric solid or haematological systemic anti-cancer treatments (e.g. drug name and doses, administration dates, and routes) (385,386). Examples of recorded systemic anti-cancer therapies include: standard cytotoxic chemotherapy; oral chemotherapy; molecular targeted therapies; biological therapy; immunotherapy; hormone therapies; chimeric antigen receptor T-cell therapy; transcatheter arterial chemoembolization; and supportive therapies prescribed alongside systemic anti-cancer therapies (e.g. steroids, bisphosphonates, antibiotics, and antiemetics) (354). SACT data records identify all incidences of systemic anti-cancer therapy for each patient during time frame stipulated at data request even if these do not apply as treatment for the primary cancer diagnosis under investigation. The individual recording of drugs utilised at the patient level sets the SACT dataset apart as prior to this, the NCRD's closest drug utilisation variable was a binary marker of utilisation/non-utilisation of chemotherapy within six months of diagnosis. The SACT dataset therefore allows differentiation of novel anti-cancer therapies from conventional cytotoxic chemotherapy as well as breaking this detail down further by drug name.

Data for patients receiving SACT treatments (entered by clinicians, nurses, pharmacists, and other healthcare providers) flows from Trusts to PHE via hospital electronic prescribing systems (354). Data of interest is abstracted from electronic prescribing and other electronic systems using system software suppliers working with local IT staff (387). As per the NCRD, Trusts vary in which electronic (prescribing) system used, though mandated monthly submission of data (by file transfer) to a secure PHE portal is standardised (354). The secure portal data repository sits behind the ENCORE interface (387). For Trusts still without electronic prescribing systems, PAS or manual systems provide the data that is included in the SACT submission (354). A registered uploader tends to be assigned to lead the monthly data submission (354). Once data flows to the secure portal, the SACT team within NCRAS at PHE hold, analyse, and release the national Trust SACT data collection (388). The SACT dataset is available to link to a range of routine care databases (354). Since the dissolution of PHE, the SACT dataset is now managed with the SACT team within the National Disease Registration Service (NDRS) at NHS Digital (389). The process of this SACT treatment data flow is depicted in Figure 3.2.

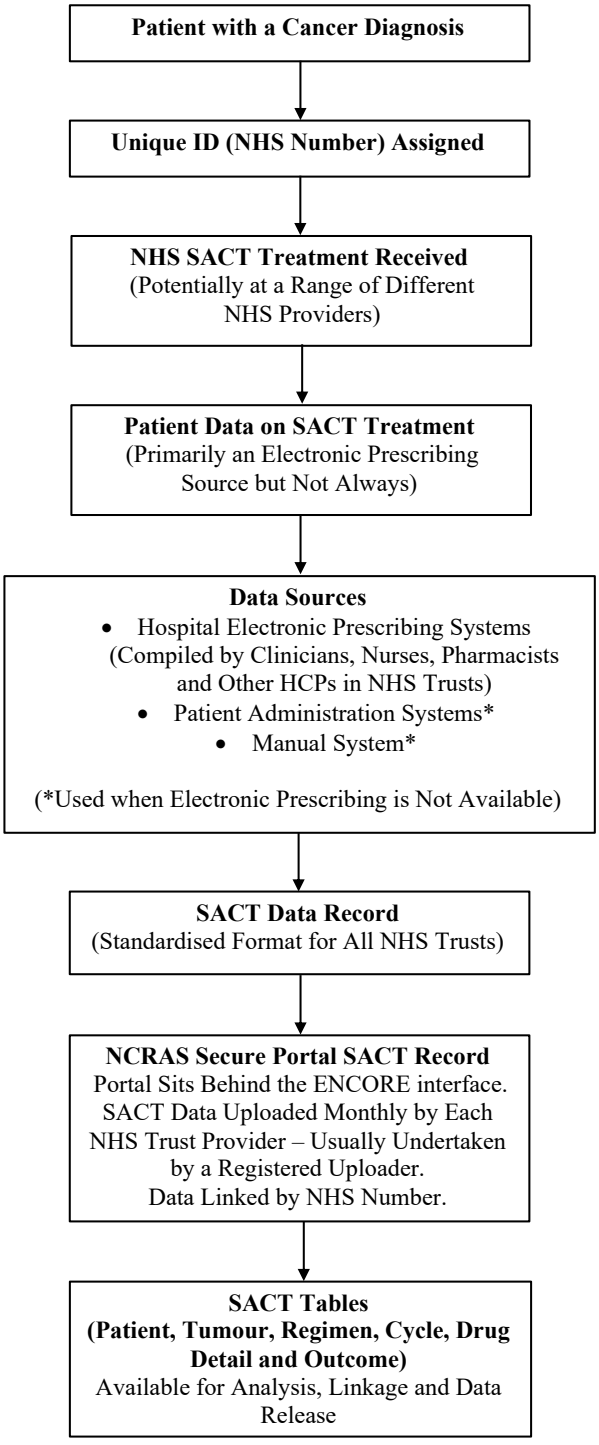


Figure 3.2 SACT data sources, data flow, and data linkage.
Please note, although all these data items may be collected, that does not mean that they are complete and/or made available to those requesting data for research purposes.
Abbreviations: ENCORE: English National Cancer Online Registration Environment HCPs: Healthcare professionals; ID: Identifier; NHS: National Health Service; NCRAS: National Cancer and Registration Analysis Service; NCRD: National Cancer Registry Database; SACT: Systemic Anti-Cancer Therapy.

SACT data are recorded in six tables: patient; tumour; regimen; cycle; drug detail; and outcome tables (354). There is some overlap of data recorded in the patient (e.g. ethnicity and age) and tumour (e.g. tumour, node, metastases - TNM) tables with the main cancer registry, thus the Office for Data Release (ODR)[5] often advise obtaining such information from the NCRD. Information recorded in the regimen (e.g. SACT programme number, drug treatment intent, and regimen grouping), cycle (number, start date, and performance status), drug (e.g. drug analysis grouping, actual dose, route of administration, and SACT administration date) and outcomes (e.g. regimen modification indicators) differentiates SACT from other NCRAS datasets (354). Table details are appropriately linked currently at the patient level, rather than the individual tumour.

Most SACT data recording has a one-to-many relationship (354). This is to reflect the real-world prescribing of these treatments, both individually and as part of a drug regimen (course of drug(s)) with repeated treatment administrations (cycles). Drug regimen is coded in SACT as two data fields: (i) analysis group which maps regimens into consistent descriptive groups to preserve treatment details (e.g. EC (epirubicin and cyclophosphamide) + docetaxel + trastuzuamb); and (ii) benchmark group which maps regimen information into consistent higher level groups (e.g. tamoxifen is mapped to "hormones"). Individual drug information is coded in the drug group data field. As patients may receive multiple drug regimens and treatment cycles, drug data recorded in SACT per patient ID can be extensive. For example, assuming a patient was to receive a full treatment course of adjuvant use intravenous paclitaxel 80 mg for 12 weeks and subcutaneous trastuzumab 600 mg for 18 cycles with no interruptions or planned changes, there would be 30 analysis and benchmark group entries (to account for mapping of both individual paclitaxel and trastuzumab drugs), along with a further 18 drug group entries for trastuzumab and 12 for paclitaxel in the patient's SACT records. However, as patients will often over time receive more than one treatment regimen, with many drug cycles, SACT data records per patient ID can quickly become extensive.

Investment in the formation of the real-world data SACT repository was driven by outcome data collection to guide new drug agent assessments (based on clinical and cost effectiveness); the data source also informs NICE funding decisions (e.g. for drugs with remaining areas of uncertainty at the end of the Cancer Drugs Fund (CDF) collection period) (392,393). This makes SACT a potentially useful tool to enable comparison of treatment delivery across

[5]ODR managed the data request at the time of the study (390). Since PHE was dissolved in 2021, NHS Digital now manages data access requests (391).

England (386). It also means that Trusts may be more likely to record expensive drugs (e.g. pembrolizumab) than cheaper, standard chemotherapies (e.g. cisplatin). The dataset has several other uses. These include: (i) publishing of SACT data in annual reports for the National Cancer Audits (e.g. National Lung Cancer Audit (NLCA)) (394) which aims to improve cancer care quality and highlight areas for improvement (354); (ii) provision of SACT data for real-world public health and epidemiological research e.g. exploration of socio-demographic patterns of care (395) and mortality of patients receiving SACT treatments (401) (see Section 3.3.3 for further SACT research previous publication details); (iii) highlighting variation between Trusts for follow up to improve cancer care delivery (including drug wastage monitoring, improving drug access, and improving clinical practice) (354,396); (iv) data tool access (e.g. CancerStats2 reporting portal gives access to interactive SACT data reports of non-identifiable data) (397); (v) dashboards of statistic metrics, including those from the SACT dataset (e.g. the CancerData platform) (398); (v) informing a simulated dataset (Simulacrum) (399) which imitates NCRAS data (including artificial SACT dataset) and has been developed to help researchers gain an insight into record-level cancer data; and (vi) use as follow up for data reported during clinical trials (especially benefits and risks of SACT treatments with longer follow up periods than the initial clinical trial may have provided for) (396).

Data collection for SACT began as a phased introduction in April 2012; by April 2014, the monthly submissions became mandatory (from e-prescribing system uploads or paper record document transcription to a secure portal maintained by PHE) and were required from all English trusts (400,401). Non–compliance with this submission results in an escalation process for investigation by NHS England and this could trigger a Care Quality Commission (CQC) inspection (385). By January 2014, 141 of 148 hospital trusts were routinely submitting mandatory data (401). Several data fields are mandated for submission and the dataset is continuously being revised. The latest version (V3.0) was launched in September 2019 and contains a number of amendments to assist the user in reporting, as well as standardising the national data received (402). Datasets can be linked by NCRAS to other cancer registry data through a patient's NHS number (385).

As per the NCRD, SACT reports only primary cancers (no secondaries), follows international data recording conventions and is subject to rigorous quality assurance checks. However, unlike the NCRD, SACT is recorded at a patient not tumour level.

3.3.3 Previous SACT Publications

The SACT dataset remains within in its infancy in terms of both data collection and analysis. However, studies are now starting to emerge with a primary focus on national SACT data (Appendix 3.1). These studies concern both patterns of care (e.g. 30 day mortality post SACT administration) (400,401) (395,403–410) and have sought to evaluate the quality of the SACT data in comparison with other linked datasets (384)(392,411,412). The utilisation of national SACT data builds upon previous real-world knowledge of SACT administration from single centre (e.g. The Christie Hospital, Manchester (413) and Royal Marsden Hospital, London (414)) and small local studies (e.g. three hospital trusts (415)). There are a few protocols referencing SACT linkage with a range of other datasets along with discussion of creation of a single research data repository to streamline population base dataset access. Examples include: the Virtual Cardio-Oncology Research Initiative (VICORI) (416); COloRECTal cancer Repository (CORECT-R) (417); and Yorkshire Cancer Research Bowel Cancer Improvement Programme (YCR BCIP) (418). Some initial publications resulting from such repository data are now emerging e.g. Taylor *et al.* (2021) (419). NCRAS have also published a SACT resource profile summary (scope, purpose, structure, quality, access, as well as the strengths and weaknesses) (354). Despite the growing interest in this new dataset, it has not yet been used to investigate associations of treatment utilisation for people with SES.

3.3.4 Ethics

Obtaining NHS ethical approval was a requirement of the ODR data request process to access NHS data at the individual patient level. Ethical approval was sought through the Integrated Research Application System (IRAS) for NHS Research Ethics Committee (REC) approval. A shorter proportionate review process undertaken because the data being applied for was already pseudonymised. Despite this, potential identification of patients within the dataset remained a significant ethical issue, especially with regards to the lung cancer data (Chapter 4) due to the rarity of some cancer sub-types and the resulting in small patient numbers. To ensure risk was negligible, the following actions were taken. Firstly, no patient identifiable data was sought and only those data fields specifically required for analysis were requested. Second, ODR only releases suitably pseudonymised data on review of the application. To assist, data fields that might be considered to increase risk of patient identification were either grouped (e.g. five year age banding) or excluded (e.g. postcode). Deprivation data was also provided in a non-identifiable format (quintile subdivisions). Third, data access was restricted to individuals in the research team who did not have additional access to any other data or information through their professional practice which when considered with the study dataset (especially in the

instances of the rarity of some lung molecular abnormalities) could potentially identify patients. For this reason, one of the supervisors of this thesis (AG), an honorary consultant oncologist specialising in lung cancer treatment, had no access to the dataset. Finally, measures were in place to ensure secure data storage at Newcastle University (e.g. data access only to named individuals, storage on a secure part of the network, analysis undertaken in offices which require a pass to access, and data not copied onto personal laptops). When combined, these measures reduced the ethical risk of being able to identify participants, to an acceptable level, a conclusion with which both ethics and ODR agreed. Favourable ethical opinion was obtained from the Proportionate Review Sub-committee of the West Midlands - Edgbaston Research Ethics Committee on October 16th 2019 (Ref 19/WM/0317).

3.3.5 ODR Application

An application to ODR at PHE was needed to access the NCRD and linked SACT data. This comprised: an ODR data request application form, research protocol, data dictionary listing the requested variables, favourable ethical approval and assurance of Newcastle University's data security and protection toolkit. At the time of the application, data was requested through PHE ODR (390). However, as of October 2021, PHE was dissolved, with NHS Digital now handling new data access requests (391).

An application for the NCRD and linked SACT records for all primary invasive tumours of the breast (ICD-10 C50.0-C50.9) [and lung (ICD-10 C34.0-C34.9) – see Chapter 4] recorded by the registry as having been diagnosed between 01/01/2012 and 31/12/2017 was requested. This period was selected to cover initiation of the dataset to the most complete calendar year of cleaned data that ODR had available for release at the time of making the data request. Both breast and lung data were requested in the same application and the data was initially supplied by ODR as a combined file. This thesis is structured to discuss each cancer separately (Chapter 3, breast and Chapter 4, lung), however, there is some overlap in the methods owing to the joint data application, which is discussed, where appropriate, throughout both Chapters.

All final data variables requested from the cancer registry were selected based on several factors. In the first instance, selection was guided by those data fields considered most likely to answer the research question at hand (e.g. data on SACT drug prescribing and IMD status). Second, discussion and advice from colleagues in NCRAS and ODR further aided decision making, especially regarding data fields with SACT and NCRD reporting overlap (e.g. staging data was requested from the NCRD as accuracy and completeness is considered superior to that

reported in SACT) or terminology confusion (e.g. requesting NCRD comorbidity data as the SACT comorbidity variable is a measure of whether a comorbidity was significant in treatment decision making as opposed to providing a number of comorbidities the patient has). Furthermore, such discussion with NCRAS also highlighted variables to avoid requesting due to issues with data completeness (e.g. SACT morphology data was poorly recorded). These conversations also informed decisions of choices between variables (e.g. different variables which may identify geographical location), since two of the major requirements of data access are that: (i) only those variables which can be clearly justified should be requested (and will be provided); and (ii) consideration needs to be given to whether combination of variables may render patients potentially identifiable. Third, variable selection was supported by the previous experiences of the PhD supervisory team with cancer registry data, and clinical expertise within the team, especially with regards to variables which are likely to influence treatment decisions, potentially confounding variables and data fields which had proved problematic to clean in previous analyses. Fourth, outputs from other cancer registry studies guided potential variable selection, especially with regards to cancer specific items (e.g. requesting receptor statuses in breast cancer). Finally, the need to ensure the data remained non-identifiable meant that specific variables (e.g. date of birth and the patients' GP practice) could not be requested. Rather, broader variable descriptors (e.g. five year age bands) were requested instead.

Several additional variables which could have been valuable to address the research question were unavailable at the time of the ODR data request. An example of such a variable was smoking status as its inclusion would have proved valuable in Chapter 4's NSCLC analyses given associations with lung cancer aetiology and low SES (105,106). Whilst this variable is now available for linkage through the NLCA, this was not an option in the NCRAS cancer registration data dictionary v3.7 (Appendix 3.2) from which the initial ODR request was made back in 2019. Similarly, whilst predictive biomarker status (HER2, ER and PR) was available for the breast cohort, no data on lung cancer mutations were available for release in 2019.

After a successful application, the fully executed data sharing contract for the project was signed by both Newcastle University and ODR on March 18th 2020. Following subsequent liaison with the NCRAS data analysts, the data was released and obtained on May 18th 2020 as five NCRD tables and one SACT table. Data on breast and lung cancer registrations was supplied combined as part of this data release. A few variables (tumour grade, diagnosis date and vital status) were missing from the initial ODR data files received and were not obtained until 29th October 2020 due to reassignment of PHE analysis to help with the COVID-19

pandemic response. A few of the requested variables were not used in the final analyses, as they were deemed superfluous to the research question once the data cleaning and analysis were undertaken (NCRD Appendix 3.3, and SACT Appendix 3.4).

Upon receiving the data, further contact was made with ODR to assist with queries. For example, there were some data code definitions for both the NCRD (MDT indicator) and SACT database (administration route, regimen modification time delay, regimen outcome summary, and regimen modification stopped early) omitted in ODR's data dictionary. Additionally, as the patient cohort was based on diagnosis date (defined according to international cancer registry conventions), it was entirely possible that some patients had drugs before their diagnosis date (either for previous or subsequent tumours or because the registry diagnosis date was after the clinical diagnosis) and hence some SACT drug regimen start dates and years of final therapy were coded as occurring outside the period of interest (2012-2017). Some date discrepancies were minor (ranging from 2000-2020) but there were also instances of dates from the 1900s, 1930s, 1940s and 1970s which were most likely input errors. ODR was contacted in the first instance to discuss for such coding issues (though explanations were not always provided).

3.3.6 Data Management: Initial ODR File Data Cleaning

Multiple NCRD data files received from ODR required amalgamating before linkage with the SACT database file prior to preparing two separate data files (one breast, one lung) for analysis. An overview of this data management process is described below (see Figure 3.3).

Step 1: Requested data fields were supplied by ODR as tables. Any data files framed at the tumour level were merged using the tumour ID data field.

Step 2: The merged file was separated by cancer type using the site of neoplasm (ICD-10 code) data field. Two files (one pertaining to only primary invasive tumours of the breast and the other, only to primary invasive tumours of the lung) then resulted.

Step 3: As this analysis was interested in treatment utilisation at the patient, not tumour level, each separate breast and lung master file was reduced to only one tumour registration per patient. This step dealt with instances where registry data can capture multiple primary tumours for the same patient (e.g. more than one breast or lung tumour)[6]. A hierarchy was created and

[6]Individual patients could be in both breast and lung ODR files if they had both a breast and lung cancer diagnosed in the study period.

repeated for analysis for both the breast and lung cancer master files, to determine which tumour record to retain for analysis (Figure 3.4). This hierarchy was as follows:

(1) The record referring to the earliest diagnosed tumour date (provided as day, month, and year) was kept.
(2) If multiple tumours diagnosed on the same day remained, the record pertaining to the most advanced stage (defined using the "best stage" variable) tumour was kept.
(3) If tumours with the same date of diagnosis and staging remained, the record with the most specific ICD-10 code was kept (i.e. ICD-10 C50.0-C50.8 for breast tumours and ICD-10 C34.0-C34.8 for lung tumours); and "non-specific" ICD-10 codes for breast, unspecified (C50.9) and bronchus or lung, unspecified (C34.9) were dropped.
(4) For any remaining tumour records with identical dates of diagnosis and staging along with a specific ICD-10 site code, the first tumour entry in the record was retained.

Step 4: Both the breast and lung cancer data files were linked with the SACT dataset (files merged using patient ID data field). This step was now only possible because the NCRD breast and lung files contained one tumour registration per patient ID.

Step 5: Upon receipt of additional ODR data files containing data fields missing in the initial data receipt, files were merged with the already prepared breast and lung master files.

This left two files (one breast and one lung cancer), cleaned to one tumour registration per patient ID ready for progression for research question specific data cleaning before statistical analysis. The master breast dataset contained 263,392 patients and the master lung dataset, 225,513 patients.

The remainder of this chapter deals with subsequent cleaning to the master breast dataset to consider a HER2+ breast population. Please refer to Chapter 4 for additional lung master file preparation.

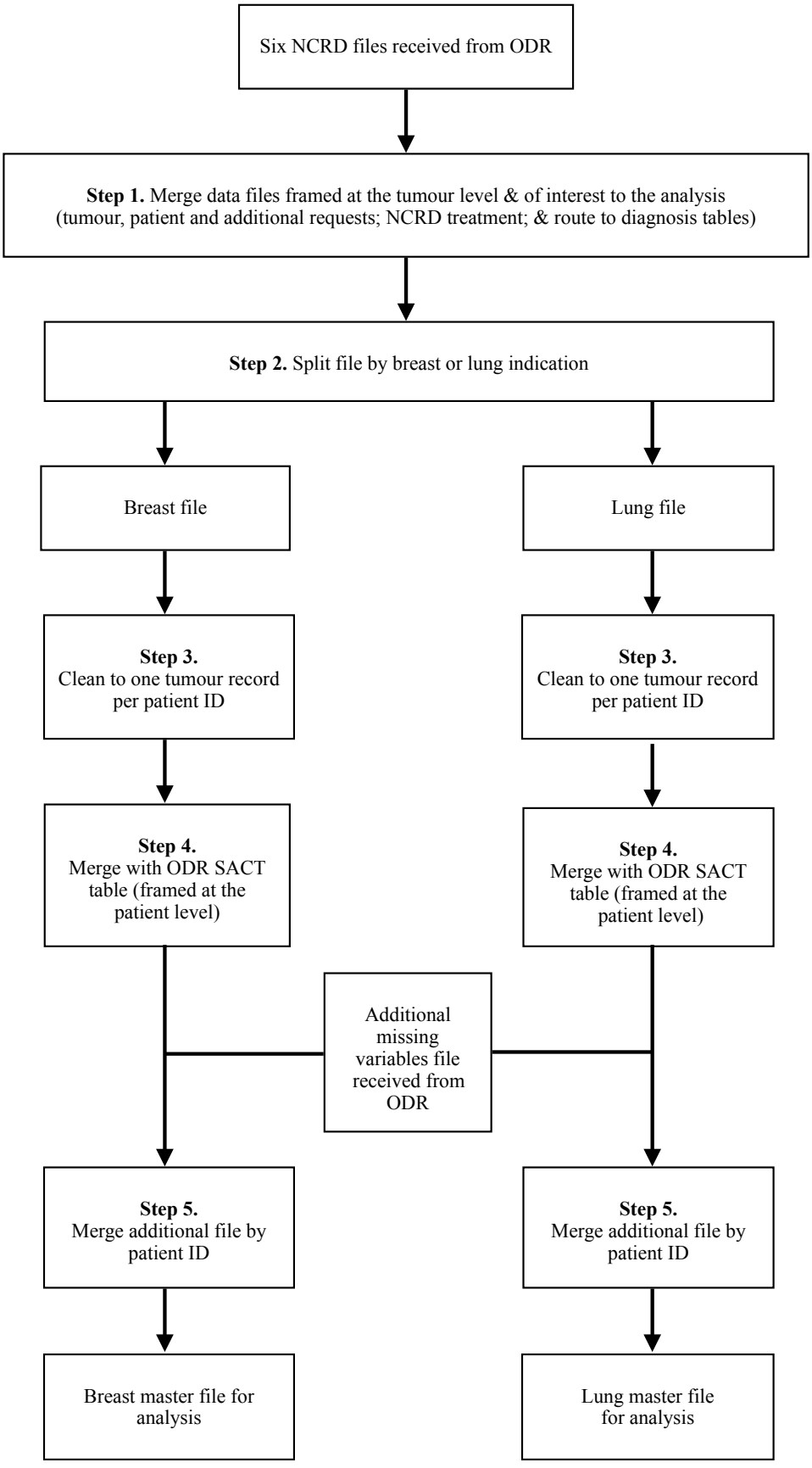


Figure 3.3 A flow diagram to describe initial data management to create two separate master files (one breast, one lung cancer). Abbreviations: ID: Identifier; ODR: Office for Data Release; NCRD: National Cancer Registry Database; SACT: Systemic Anti-Cancer Therapy Database.

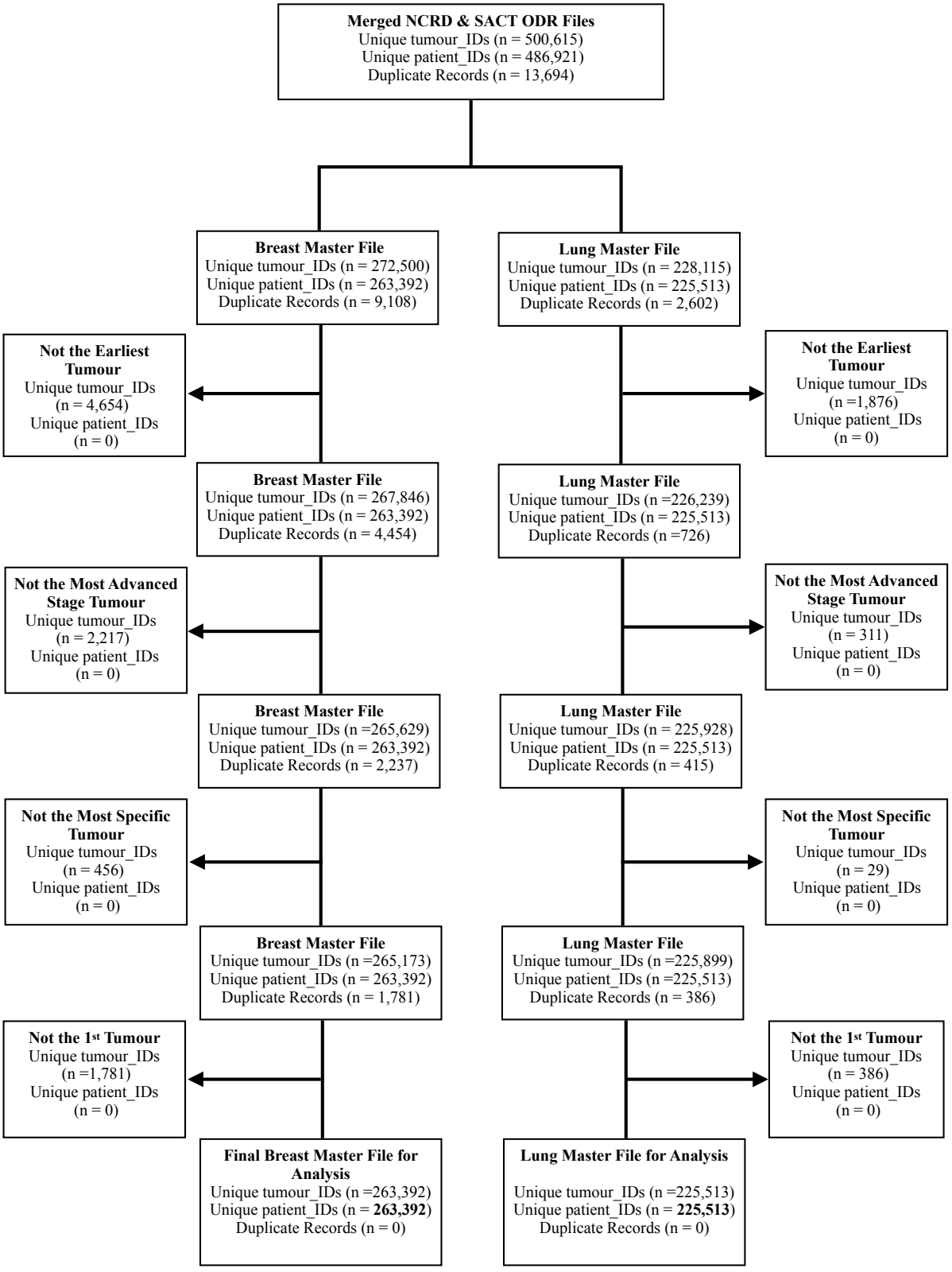


Figure 3.4 A flow chart to show how the master merged NCRD & SACT ODR file was reduced to one tumour registration per patient ID for both breast and lung cancers. Abbreviations: ID: Identifier; ODR: Office for Data Release; NCRD: National Cancer Registry Database: SACT: Systemic Anti-Cancer Therapy Database.

3.3.7 Data Management: Breast File Preparation for Analysis

The cohort of interest for the breast analysis was women with HER2+ tumours as this is the group of patients for whom trastuzumab is a treatment option. To undertake this analysis, several inclusion and exclusion criteria were applied to the breast master file to provide a denominator population of HER2+ breast cancer patients for analysis. These are depicted in Figure 3.5 (page 109) and are described further below.

As male breast cancer is rare, males (n = 1,815) were excluded from the analysis, and this left a total of 261,577 female patients diagnosed with a breast cancer between 2012-2017. The breast cohort was then refined using receptor status, with the denominator population being defined as all cases of a HER2 positive breast cancer. The "HER2 status" NCRD variable was used to refine the dataset. HER2 status was obtained from pathology results[7] (patient's immunohistochemical status where performed by clinical teams) (421). In-situ hybridisation tests were only required and reported if the initial HER2 status was deemed "borderline" (421). The variable defines categorisation of HER2 status and is used in this instance as a proxy measure for receipt of HER2 biomarker testing.

For HER2 status, test results are recorded in the registry as follows:

- N – Negative
- P – Positive
- X – Not performed
- B – Borderline
- Pm – The patient has one positive test but may have had a different result in other tests (one record as P but at least one record with something else).
- (.) – Missing

For HER2 status, tumours were classified as follows:

- P, B and Pm were defined as "positive"
- N as "negative"
- X and (.) as "missing"

"Positive" and "Negative" classifications therefore defined HER2 status.

[7]Pathology systems are key routine data sources for the NCRD with tumour characteristics being mandatory items for reporting (382,420).

Of the 208,420 cases with a defined HER2 status, 40,278 (19.3%) were HER2+. Tumours classified as stage 0 were also excluded due to small numbers (n = 99) and expert clinical opinion that systemic treatment in such staging would be uncommon (only 18 of these patients had a SACT record). At the end of this process, this left a denominator, analytical population of 40,179 patients.

Of the denominator population, 14,580 patients had no linked SACT record. For those patients who did have a linked SACT record (n = 25,599), further data cleaning was required. This was because there can be multiple drug entries in SACT per patient ID depending on how many and how often the patient received each SACT treatment. As this analysis was only concerned with whether a patient utilised trastuzumab and not the number of times they received trastuzumab, the data file was cleaned to leave only one SACT entry per patient ID. To assist with this and to increase the likelihood that SACT treatment were used for the primary invasive HER2+ breast cancer of interest, a time frame restriction was applied. Consideration was restricted to SACT administered within 56 days prior to and one year after the date of diagnosis of the incident breast cancer. This period was selected to reflect the treatment timeline for a "standard" HER2+ patient receiving surgery, chemotherapy (adjuvant or neoadjuvant) and anti-HER2 therapy. A limit of 56 days prior to diagnosis was chosen to capture discrepancies in how the NCRD and SACT record "data of incidence" (for example, where SACT is started for the cancer of interest prior to official date of diagnosis recording in the NCRD - which follows a standard algorithm). At the end of this data cleaning process there were 21,881 patients with a SACT record entry within the desired time frame out of a total analytical cohort of 40,179 patients.

3.3.8 Missing and Unknown Data

Missing data in the registry was to be expected and this is problematic in clinical research as it can lead to bias and loss of information (422). Missing data can be complex, and methods are needed to deal with its presence. Experience in the supervisory team with cancer registry data meant that there was a recognition of the fact that registry data is often not missing at random; thus, simply dropping cases with missing data, or restricting analyses to a "complete cohort" is likely to be inappropriate. In this analysis, missing data was handled first by reporting the number of missing values for each variable of interest and including this information on the analytical cohort flow diagram (where appropriate). Decisions were also made to avoid using data variables with substantial missing data (e.g. performance status, which had over 50% missing data). Cases with missing data were included to retain the population-basis of the

analysis and to minimise bias. A separate "missing" category was included where possible when missing data was high, or such information was grouped (e.g. missing/unknown/other) when missing numbers were smaller (Table 3.1).

Over a third of the analytical cohort had no SACT record. A SACT record is only generated if a patient has systemic drug therapy. Thus, having a null value for this variable does not mean that such data was "missing" as there may be a valid clinical reason (e.g. drug not offered) or otherwise (e.g. patient choice) as to why no drug treatment record exists. The demographics and clinical characteristics of patients with and without a SACT record were compared. However, no clear differences were observed. There is also no plausible hypothesis as to why a patient with a low SES would be less likely to have had a documented SACT record than a patient with high SES (i.e. that data for a treated patient of low SES would be more often missing/unrecorded than data for a treated patient of high SES). The decision was therefore taken to retain the "no SACT record" unknown treatment data as part of the analytical cohort, and to assume that these patients did not receive treatment. Finally, limitations regarding including missing data are reported throughout this thesis.

There are other missing data handling strategies not considered in this analysis as their use is not common with this type of data e.g. multiple imputation. This method requires creation of dataset copies, with missing values replaced by an imputed value determined from the predictive distribution in the observed data (422). Models can then be fitted to imputed datasets and averaged together to provide overall estimated associations (422). Whilst multiple imputation can be useful for improving validity, it is only one technique for handling missing data, and does have limitations (e.g. assuming that data are normally distributed, the missing at random assumption plausibility, issues with instances where data are missing not at random, and algorithm computation issues to name a few) (422). Given this, as well as the fact that the analysis included a mix of both negligible (e.g. grade) and substantial (e.g. SACT record) "missing" data, the decision was taken instead to report the observed data – and discuss the extent of the missing data and its implications. This method of missing data handling is also considered acceptable (423).

Table 3.1 Missing, unknown, and other data in the HER2+ breast cancer dataset. Abbreviations: ER: Oestrogen receptor status; MDT: Multidisciplinary team; SACT: Systemic anti-cancer therapy.

Variable	**Missing/ Unknown Data**	**Data Handling Strategy**
No SACT Record	14,580 (36.29)	Included in the analytical cohort as a SACT record is only generated if a patient has systemic anti-cancer drug treatment. Characteristics compared to those in the group with a SACT record. Highlighted in limitations as high proportion of unknown data.
Ethnicity	1,813 (4.51)	Labelled as "missing/unknown". Included in multivariable analyses.
Stage	3,194 (7.95)	Labelled as "unknown". Included in multivariable analyses. Unlikely missing as random – likely mainly comprises older patients with more advanced disease (and with other comorbidities) who were not sent for staging.
Grade	847 (2.11)	Labelled as "other". Included in multivariable analyses.
ER Status	6,107 (15.20)	Labelled as 'unknown". Included in multivariable analyses. Highlighted as a limitation.
Discussed at MDT	6,159 (15.33)	Labelled as "missing". Included in multivariable analyses. Highlighted as a limitation.
Performance Status	23,218 (57.79)	Unable to use in analysis as substantial missing data (over 50% of patients) and only available for those with a SACT record.

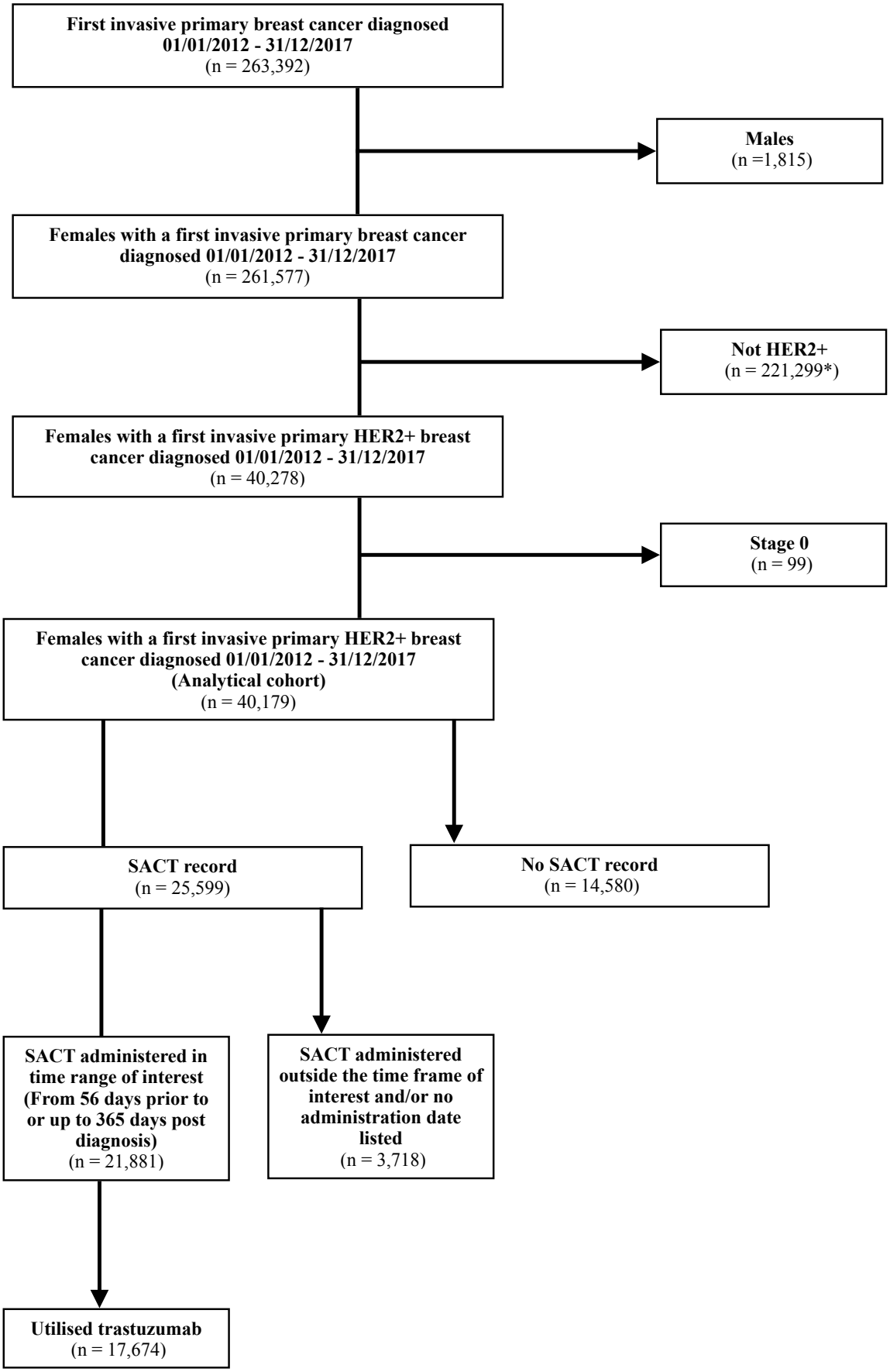


Figure 3.5 Flow diagram depicting the analytical cohort. Abbreviations: HER2+: Human epidermal growth factor receptor 2 positive; SACT: Systemic anti-cancer therapy. *Includes missing values (n = 53,157).

3.4 Outcome Measures

Primary Analysis - Trastuzumab Utilisation: The primary focus of this Chapter was to determine whether a patient utilised trastuzumab. The rationale for focusing on trastuzumab was because this treatment is an established targeted therapy and whilst there are other novel anti-cancer therapies utilised in HER2+ breast cancer treatment (e.g. pertuzumab) recorded in the SACT dataset, these treatments were uncommon in the time window of the study.

As SACT data does not contain information on drug indication, reference to trastuzumab within SACT data needed defining.

Trastuzumab utilisation was defined as follows:

- A record of trastuzumab in any of the drug, benchmark, or analysis group data fields.
- A record of trastuzumab, either alone or in combination with other treatments e.g. trastuzumab and pertuzumab.
- A reference to a clinical trial where it was clear that all patients would have received trastuzumab, regardless of the trial arm. Information from trial databases such as CRUK clinical trials finder (424) or clinical trials.gov (425) (if the trial was still active), trial protocols or from guidance from two oncologists (AG and NC)[8] was used to assist with discerning this.

Trastuzumab utilisation was defined by working backwards using all SACT data recorded in the denominator population records. Several steps helped minimise misclassification of trastuzumab containing regimens. First, NICE guidelines (97,426), the BNF (246), Summary of Product Characteristics (SPC) (427) and EMA website information (428), local Northern Cancer Alliance protocols (429) and Newcastle-upon-Tyne Hospitals NHS Foundation Trust regimen upload list were checked. Second, the American Society of Clinical Oncology (ASCO) (203) and European Society for Medical Oncology (ESMO) guidelines (430) were also reviewed for completeness to capture any further trastuzumab containing regimens. Third, two oncologists (AG & NC) exercised clinical judgement, excluding unlikely drug regimens e.g. trastuzumab use in gastric cancer. Any treatments with no HER2+ breast cancer indication were excluded.

[8]AG refers to Dr. Alastair Greystoke and NC to Nicola Cresti.

Of the 323 analysis groups, 306 benchmark groups and 141 drug groups supplied as part of the ODR data request, only 39 analysis groups, 36 benchmark groups and 9 drug groups were included as meeting the definition for a reference to "utilisation of trastuzumab" (Appendix 3.5). For each drug/SACT entry per patient ID, a binary variable was created (utilisation or non-utilisation of trastuzumab using Appendix 3.5). Duplicate references to trastuzumab for an individual patient in SACT data were then dropped (the first instance of each drug record per patient was kept). This left only one drug record per patient ID (n = 40,179).

Secondary Analysis - Breast Cancer Directed Surgery Utilisation: As a sense check to the data and to compare whether patterns of utilisation differed by treatment type (novel or conventional), utilisation of conventional therapies was also explored given the historically documented reduced utilisation with low SES. Surgery receipt within six months of diagnosis was supplied as a binary (Y/N) indicator flag by ODR.

Secondary Analysis - Chemotherapy Utilisation: As per the surgery analysis, exploring chemotherapy utilisation provided a further sense check to the data and enabled comparison of whether patterns of utilisation varied by treatment type (conventional or novel). Chemotherapy receipt within six months of diagnosis was supplied as a binary (Y/N) indicator flag by ODR.

3.4.1 Main Explanatory Measure: Deprivation

Deprivation was used as a relative measure of SES. In the NCRD, the measure of deprivation recorded is the IMD (using only the income domain). IMD in NCRD is grouped into quintiles and is the measure of relative deprivation for small areas (or neighbourhoods) in England (383). The measure ranks every lower layer super output area (LSOA) from the most current census (of which there were 32,482 in 2010 and 32,844 in 2015) from the most deprived (IMD 5) to the least deprived (IMD1) (32,844) (383,431). LSOAs contain an average of 1,500 people or around 650 households (383).

IMD is a composite measure of seven domains of deprivation with different weightings (income 22.5%; employment 22.5%; education, skills and training 13.5%; health deprivation and disability 13.5%; crime 9.3%; barriers to housing and services 9.3%; and living environment 9.3%) (383,432). NCRAS use postcodes of the residence at the time of diagnosis to assign each person a level of deprivation (IMD) for reporting in the registry. English cancer registries now only use the income domain of the IMD measure. Specifically, this measures the proportion of the population experiencing deprivation relating to low income (taken as both those out of work

as well as those that are in work but who have low earning and therefore who satisfy the respective means tests) (433).

IMD income domain data was supplied by ODR as both the IMD 2010 and IMD 2015 variables. As the dataset covered the period 2012 to 2017, for the diagnosis year 2012, the IMD 2010 variable was used to define deprivation and for the diagnosis years 2013, 2014, 2015, 2016 and 2017, IMD 2015 was used. This was to ensure that the deprivation status closest represented the LSOA at the time of data collection.

3.4.2 Other Explanatory Measures[9]

Age: Diagnosis age for breast cancer patients was organised into six groups (<40, 40-49, 50-59, 60-69, 70-79 and 80+).

Year of Diagnosis: Period of diagnosis was determined by diagnosis year. This variable was considered a covariate of interest given that changes in drug availability can occur over time, so it was important to capture this variable in statistical models. Year was also an important factor for consideration in sensitivity analyses given the changes in SACT recording following mandatory SACT uploads for all hospital trusts as of April 2014.

Ethnicity: Ethnicity data was supplied by ODR in 21 categories. Ethnicity was grouped as: white; other ethnic group (Asian/British Asian, Black/African/Caribbean/Black British, mixed/multiple ethnic groups and other ethnic groups); and missing/unknown (refers to missing and unknown ethnicity classifications). The decision was taken to group ethnicity in this way because the white ethnic group formed most of the dataset (n = 35,276; 87.8%) and other individual ethnic groups were rare.

Rural/Urban Indicator: Rural/urban indicator as supplied by ODR is a standard classification tool (434). The decision was taken to collapse some of these standard categories in instances where case numbers were small (for example, in both the sparse and non-sparse setting, “hamlets and isolated dwellings” was merged with “villages”). The final categories used were: rural village (including in a sparse setting); hamlet and isolated dwellings (including in a sparse setting); rural town and fringe (including in a sparse setting); urban city and town (including in a sparse setting); urban conurbation (minor and major).

[9]Not all explanatory measures of potential interest were available at the time of data request. This limitation is discussed further in the discussion (see Section 3.7.3).

Government Region: Government region was provided categorised as: North West; North East; West Midlands and the Humber; East Midlands; East of England; South East; South West; and London. These government office regions were established in 1994 and became the primary classification in regional statistics by 1996 (435).

Stage: Stage was grouped as stage I, II, III, IV and unknown (missing or unstageable tumours) using the "stage_best" variable. Staging subdivisions were combined (e.g. 1, 1A, 1A1, 1B, 1C were classified as stage I).

Grade: Grade was grouped as the following: well differentiated; moderately differentiated; poorly differentiated; and other (anaplastic, undetermined, and missing).

Multiple Tumours: This NCRD variable, called "big tumour count", provides an indication for whether a patient has one or more than one primary tumour (ICD-10 C00-C97 but excluding C44 referencing non-melanoma skin cancer as registration of this cancer is generally accepted to be incomplete). "Big tumour count" was supplied by ODR as a number ranging from 1 to 5. This was simplified to classification by two groups (one primary tumour or more than one primary tumour).

Oestrogen Receptor Status: ER status was retained for the analysis as this receptor status is considered important in a HER2+ cancer diagnosis in so much that it increases the range of treatments available; the prognosis for patients with ER+ and HER2+ breast cancer is usually superior compared to ER- and HER2+ (436). ER status was classified as follows: positive (positive, borderline, patient has one positive test even if they have different result in other tests); negative (negative); and unknown (unknown, not performed, missing).

Progesterone receptor status testing is not always routinely performed and is arguably less important in treatment decision making for a HER2+ breast cancer patient setting, so was omitted from analyses.

Comorbidities: ODR provides a data field recording the number of comorbidities a patient has. This field is created by NCRAS staff using Hospital Episode Statistics (HES) records for the time period between 6 and 78 months prior to the cancer diagnosis. This variable was grouped as 0, 1-2, or 3+ comorbidities in the main cancer registry and reports comorbidities diagnosed between 78 to 6 months prior to diagnosis. The ODR variable is derived from the Charlson

Comorbidity Index Score (CCM) (437). This was developed as a method of predicting mortality by classifying or weighing comorbidities to measure the burden of disease (437). It is derived from multiple sources of hospital administration data relating to a series of defined conditions (379). This means that the CCM has limitations as a comorbidity measurement based on hospital admission conditions or hospital stay complications likely under ascertains true levels of comorbidities, especially as those conditions diagnosed and treated in a primary care setting are not included (104,438). When calculating the comorbidity score, ODR does not count the primary invasive cancer. However, if a patient with a primary invasive breast cancer went on to develop another cancer, this tumour would be given a Charlson score of 2 to account for the prior breast diagnosis. Cancer registry data are used to generate comorbidity scores for primary cancers following the methodology outlined in Quan *et al.* (2005) (439). HES diagnostic codes for inpatient admissions are used to generate a score for all other conditions.

Discussion at MDT: This variable was classified as: yes; no; or missing.

3.5 Statistical Analysis

The purpose of the primary analysis was to determine the association of SES with trastuzumab utilisation. A secondary analysis then determined the association of SES with conventional (breast cancer directed surgery and chemotherapy) treatment utilisation.

Baseline demographic and clinical characteristics were reported for the whole cohort (n = 40,179), those without a SACT record and those without a HER2+ status. For each treatment type (trastuzumab, breast cancer directed surgery and chemotherapy) descriptive statistics (number and percentage utilising) were computed by all independent variables of interest. Chi-square tests determined if there was an association between the demographic/clinical variable and utilisation of treatment. Additional chi-square tests also aided determining associations between deprivation and all other explanatory variables.

Logistic regression models were developed to determine associations between SES and (i) trastuzumab utilisation; (ii) breast cancer directed surgery utilisation; and (iii) chemotherapy utilisation, with and without adjustment for covariates. Any explanatory clinical and demographic variables were considered for inclusion in the multivariable model if these variables were significant in the univariable analyses (likelihood ratio test (LRT) $p \leq 0.05$). However, SES as the primary variable of interest, was forced into all models. Models were reduced so that they only contained variables that remained statistically significant (LRT $p \leq$

0.05) in the presence of other variables. Bayesian Information Criterion (BIC) and Akaike Information Criterion (AIC) were used to inform choices between models and goodness of fit was determined using the Hosmer and Lemeshow χ^2 test (440). All model mean variance inflation factors were lower than 10. In these analyses, deprivation was fitted as a categorical variable. In the final multivariable model, a test for linear trend across deprivation categories was also conducted.

When preparing models for both breast cancer directed surgery and chemotherapy utilisation, interactions were present. To deal with this in the surgical model, an age-stage interaction variable was created and grouped as follows: early/late stage and ages <50, 50-59, 60-69, 70+. For the chemotherapy analysis, the cohort was stratified by surgery receipt. Age in this analysis was classified as <60 or ≥60 years old. The rationale for splitting the cohort in this way was because chemotherapy is used in two ways (adjuvant treatment in women who have had surgery and in women who have not had surgery - generally these tend to present at a more advanced stage). This age cut off was used because it was deemed important to incorporate some measure of age into the models as opposed to dropping this variable. As there is a difference between pre- and post-menopausal breast cancer treatment, a <60, ≥60 year age limit seemed an appropriate compromise to provide a working model.

The primary analysis final trastuzumab multivariable model was adjusted for the following covariates: age; diagnosis year; ethnicity; rural/urban categorisation; government region; stage; grade; ER status; comorbidities; and whether discussed at MDT. The breast cancer directed surgery final multivariable model was adjusted for: age stage interaction; diagnosis year; ethnicity; rural/urban categorisation; grade; ER status; comorbidities; and whether discussed at MDT. Finally, the final multivariable chemotherapy model was adjusted for: age; diagnosis year; ethnicity; rural/urban categorisation; government region; stage; grade; ER status; comorbidities; and whether discussed at MDT. Unadjusted and adjusted ORs are reported with 95% CI and p values for all models.

Several sensitivity analyses were performed. These restricted the denominator population by: (i) a diagnosis date post April 2014 (all models); (ii) a HER2 status positive classification only (all models); (iii) age <60, ≥60 years old (trastuzumab and surgery models); and (iv) SACT record in time range only (trastuzumab model). Time period sensitivity analyses were conducted to reflect the fact that SACT data recording prior to April 2014 is known to be poorer. Restricting the cohort to a refined HER2 definition enabled tighter classification of the

definition of a positive result. Age restricted analyses were undertaken to permit comparison with the age-stratified chemotherapy models and to understand if there might be differential effects of SES on trastuzumab and surgery utilisation in younger and older women. Finally, as a significant proportion of the denominator population had no linked SACT record, the final multivariable model was re-run in a restricted population (those patients with a SACT record in the time frame) for trastuzumab utilisation only.

All statistical analyses were conducted using STATA version 16.1 (StataCrop, College Station, Texas (TX)).

3.6 Results

3.6.1 Cohort Characteristics

All Patients: There were 40,179 patients diagnosed with first, invasive primary HER2+ breast cancer between 01/01/2012 - 31/12/2017. Of these patients 87.8% were white, 4.8.2% were aged between 50-69 years old, 45.5% resided in urban cities and towns and 81.9% had no comorbidities (Table 3.2). Clinically, these women tended to have stage I or II tumours (74.1%), be graded as moderate or poorly differentiated (92.5%) with an ER+ receptor status (63.4%). The percentage of patients' residing in each deprivation quintile was as follows: 23.1% for IMD 1 (least deprived), 22.3% for IMD 2, 20.2% for IMD 3, 18.4% for IMD 4 and 16.0% for IMD 5 (most deprived). The number of primary invasive HER2+ breast cancer cases diagnosed each year increased over the time period the data covered (11.7% of the total HER2+ breast cancer were diagnosed in 2012 and 21.5% in 2017) (Table 3.2).

Table 3.2 Demographic and clinical characteristics of women with HER2+ breast cancer diagnosed between 01/01/2012 - 31/12/2017 (n = 40,179)

Characteristic	Number (%)
IMD[1]	
1 (Least Deprived)	9,272 (23.08)
2	8,967 (22.32)
3	8,127 (20.23)
4	7,375 (18.36)
5 (Most Deprived)	6,438 (16.02)
Age at Diagnosis (Years)	
<40	2,616 (6.51)
40 – 49	7,154 (17.81)
50 – 59	9,778 (24.34)
60 – 69	9,594 (23.88)
70 – 79	6,475 (16.12)
80+	4,562 (11.35)
Diagnosis Year	
2012	4,716 (11.74)
2013	5,745 (14.30)
2014	6,219 (15.48)
2015	6,980 (17.37)
2016	7,902 (19.67)
2017	8,617 (21.45)
Ethnicity	
White	35,276 (87.80)
Other Ethnic Group[2]	3,090 (7.69)
Missing/unknown[3]	1,813 (4.51)
Rural/Urban Indicator	
Rural Village, Hamlet & Isolated Dwellings	4,374 (10.89)
Rural Town & Fringe	4,295 (10.69)
Urban City & Town	18,265 (45.46)
Urban Conurbation	13,245 (32.96)
Government Region	
North West	6,309 (15.70)
North East	2,532 (6.30)
West Midlands	4,302 (10.71)
Yorkshire & the Humber	3,560 (8.86)
East Midlands	3,655 (9.10)
East of England	5,188 (12.91)
South East	6,483 (16.14)
South West	4,349 (10.82)
London	3,801 (9.46)

[1]For diagnosis year 2012, IMD_2010 was used and for diagnosis years 2013-2017, IMD_2015 was used.
[2]Other refers to Asian/British Asian, Black/African/Caribbean/Black British, mixed/multiple ethnic groups and other ethnic groups.
[3]Missing/unknown ethnicity refers to unknown and missing ethnicity classifications.

Table 3.2 Continued

Characteristic	Number (%)
Stage	
I	13,094 (32.59)
II	16,663 (41.47)
III	4,859 (12.09)
IV	2,369 (5.90)
Unknown[4]	3,194 (7.95)
Grade	
Well Differentiated (Low Grade)	2,177 (5.42)
Moderately Differentiated	17,391 (43.28)
Poorly Differentiated	19,764 (49.19)
Other[5]	847 (2.11)
Multiple Tumours	
1	35,303 (87.86)
>1	4,876 (12.14)
ER Status	
Positive	25,471 (63.39)
Negative	8,601 (21.41)
Unknown	6,107 (15.20)
Number of Comorbidities (Between 78 to 6 Months Prior to Diagnosis)[6]	
0	32,899 (81.88)
1 – 2	5,902 (14.69)
3+	1,378 (3.43)
Discussed at MDT	
Yes	27,783 (69.15)
No	6,237 (15.52)
Missing	6,159 (15.33)
SACT Record	
Yes (In Time Range)[7]	21,881 (54.46)
Yes (Not in Time Range)[7]	3,718 (9.25)
No	14,580 (36.29)
Treatment	
Utilised Chemotherapy	24,444 (60.84)
Utilised Surgery	30,124 (74.97)
Utilised Trastuzumab	17,674 (43.99)
Utilised Trastuzumab & Surgery	13,369 (33.27)
Utilised Trastuzumab but not Surgery	4,305 (10.71)
Utilised Trastuzumab & Chemotherapy	17,461 (43.46)
Utilised Trastuzumab, Surgery & Chemotherapy	13,224 (32.91)

[4]Unknown staging refers to missing and unstageable tumours.
[5]Other refers to undifferentiated or anaplastic, undetermined, and missing tumour grades.
[6]Refers to Charlson Comorbidity Index.
[7]Refers to between 56 days prior to or up to 365 days post diagnosis.
Abbreviations: ER: Oestrogen receptor status; IMD: Index of Multiple Deprivation; MDT: Multidisciplinary Team; SACT: Systemic Anti-Cancer Therapy.

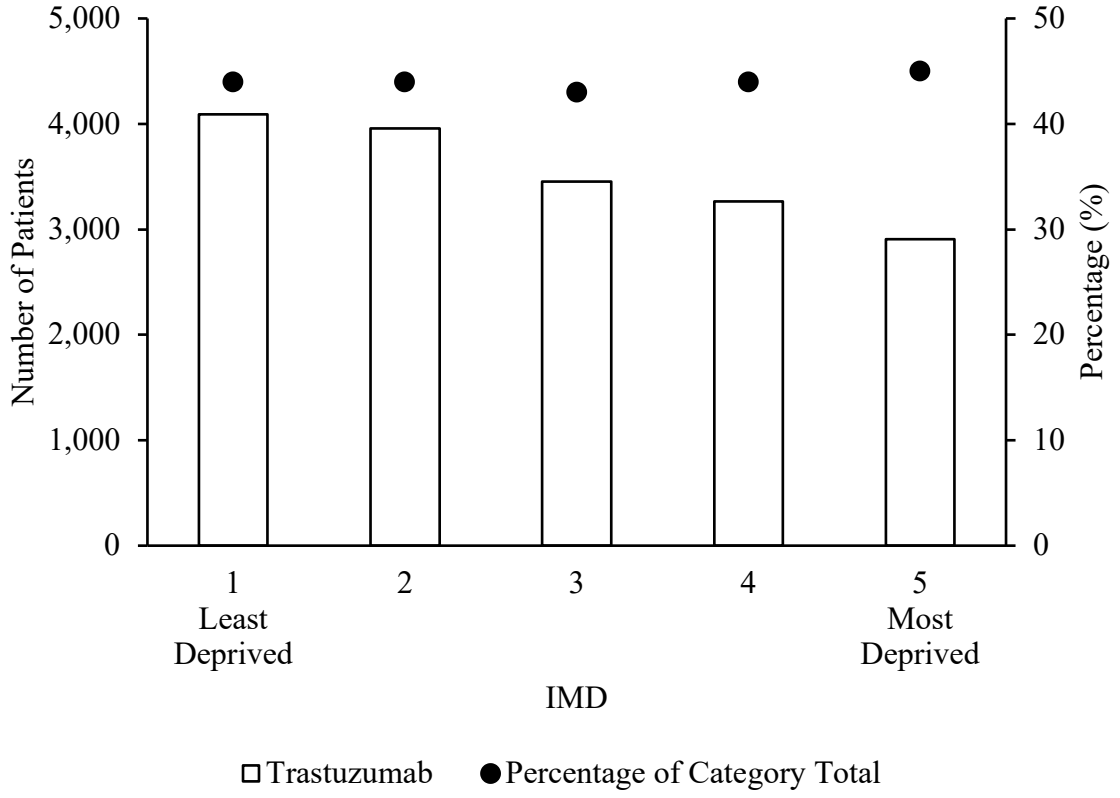
SACT Record in the Time Range: Just over a third of women (36.3%; n = 14,580) in the dataset had no SACT record. For those women who did have a SACT record within this time frame (n = 21,881), similar trends in demographic and clinical characteristics when compared to full cohort (n = 40,179) were observed (Appendix 3.6). The number of patients with a linked SACT record did not change much during the time frame of the analysis (in 2012, 61.2% of total records had a SACT record and 60.3% in 2017).

Demographic and Clinical Characteristics by Deprivation Category: Demographic and clinical characteristics by deprivation category are shown in Appendix 3.7. Only the variables diagnosis year and multiple tumours had non-significant p values. For women with the highest deprivation category (IMD 5), the proportion of non-white ethnicity, a comorbidity score of greater than one, residence in an urban conurbation and diagnosis at an age <40 were all higher when compared to those women with the lowest deprivation status.

3.6.2 Primary Analysis: Trastuzumab Utilisation

Descriptive Statistics: 44% of women (n = 17,674) utilised trastuzumab. Utilisation was consistent across all deprivation quintiles: 44% for IMD 1 (least deprived); 44% for IMD 2; 43% for IMD 3; 44% for IMD 4; and 45% for IMD 5 (most deprived) (Figure 3.6a). Utilisation of trastuzumab increased over the time; from 33.2% of patients diagnosed in 2012 to 43.9% of patients diagnosed in 2017 (Figure 3.6b). 72.8% of women utilising trastuzumab had stage I or stage II cancers (Figure 3.6c). Of women utilising trastuzumab, 99% also received chemotherapy, though slightly fewer utilised trastuzumab and underwent breast surgery (76%) or received both chemotherapy and surgery in addition to trastuzumab (75%) (Figure 3.6d).

a)



b)

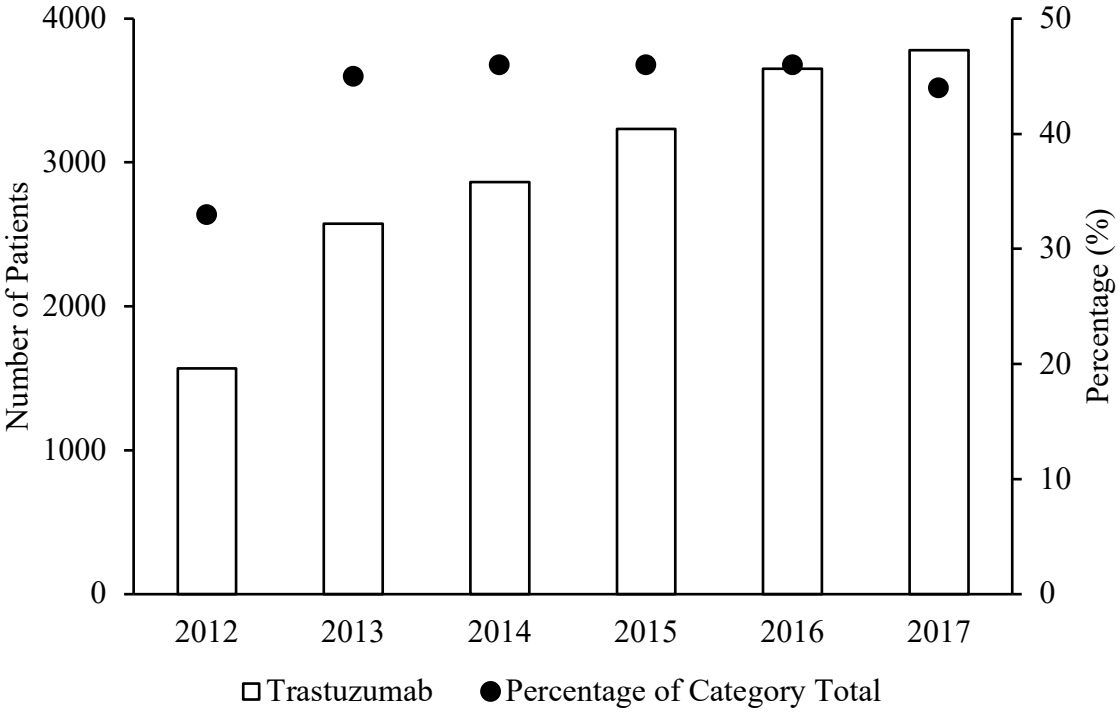
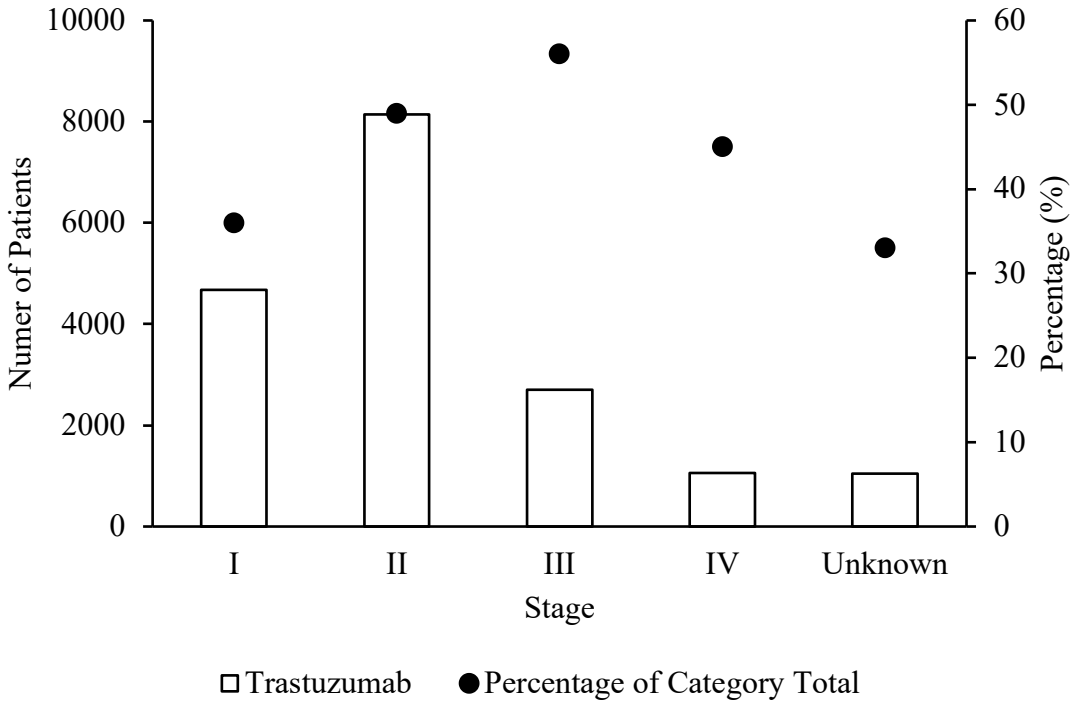


Figure 3.6a Trastuzumab utilisation by **a)** each quintile of deprivation as measured by IMD (IMD 1 = least deprived) and **b)** by diagnosis year IMD: Index of Multiple Deprivation.

c)



d)

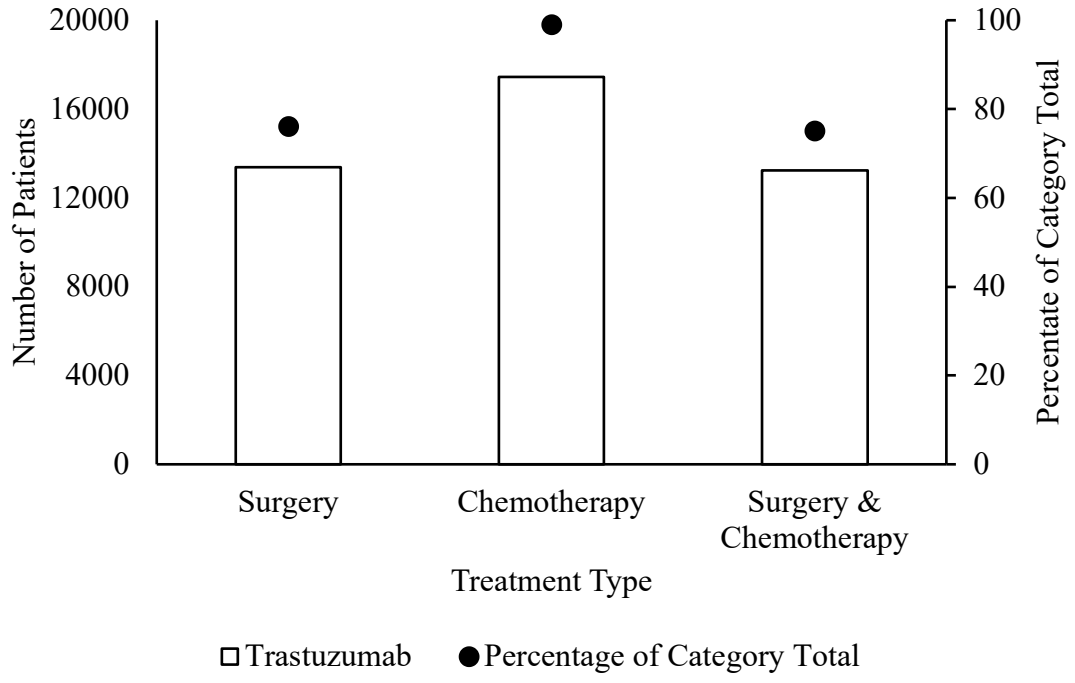


Figure 3.6 Trastuzumab utilisation by **c)** stage and **d)** in combination with conventional therapies.

Multivariable Model Deprivation Associations: Utilisation of trastuzumab was similar regardless of the level of deprivation: 44.1% (IMD 1 least deprived); 44.2% (IMD 2); 42.5% (IMD 3); 44.3% (IMD 4); and 45.1% (IMD 5 most deprived). However, after adjustment for confounders in the final multivariable model, there was a significant association between trastuzumab utilisation and deprivation (LRT $p \leq 0.05$). There was also a trend across deprivation categories for reduced trastuzumab utilisation with increasing deprivation status and the test for linear trend was found to be significant ($p \leq 0.05$). This pattern of decreasing odds of utilisation with increasing deprivation was more pronounced from IMD 3 through to IMD 5 where ORs were also less than 1. Here the odds of utilising trastuzumab was 7-8% lower than women resident in IMD 1 (Figure 3.7). Women living in the most deprived areas were found to be less likely to utilise trastuzumab compared with women living in the least deprived areas (IMD 5 vs IMD 1; multivariable odds ratio (mvOR) 0.92, 95% CI 0.85, 0.99) (Table 3.3).

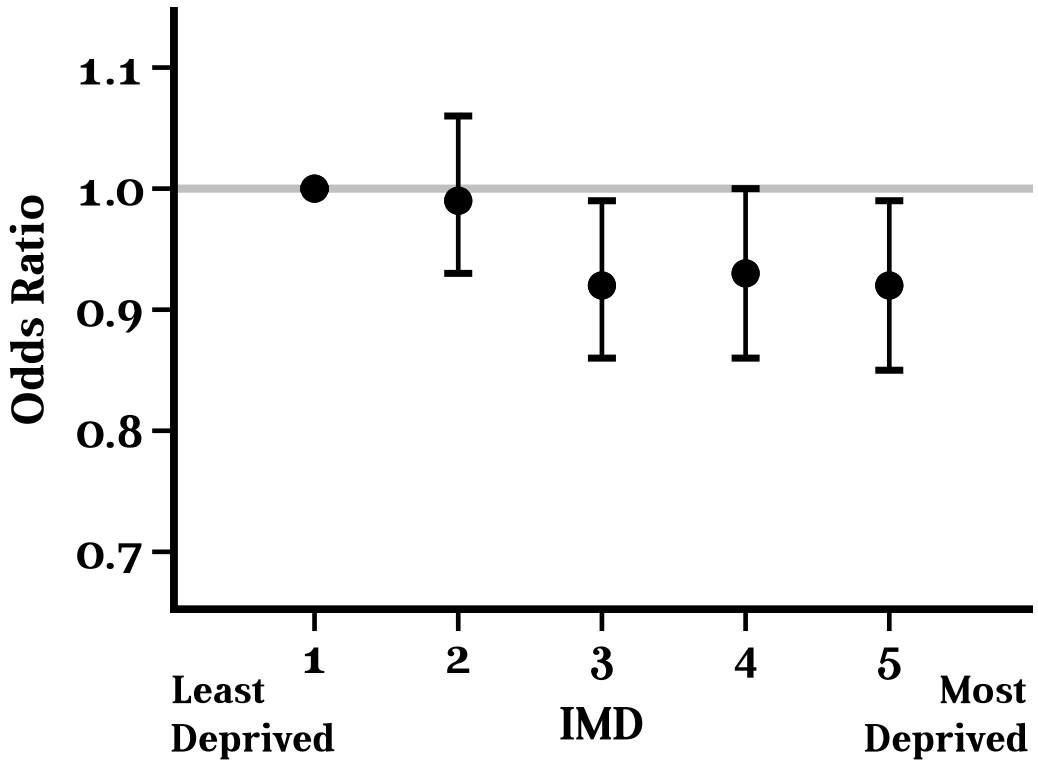


Figure 3.7 Trastuzumab utilisation by deprivation (IMD 1 least deprived; IMD 5 most deprived). Trastuzumab multivariable model adjusted for: age, diagnosis year, ethnicity, rural/urban categorisation, government region, stage, grade, ER status, comorbidities, and whether discussed at MDT. Abbreviations: ER: Oestrogen receptor; IMD: Index of multiple deprivation; MDT: Multidisciplinary team.

Sensitivity Analyses: When restricting the cohort to breast cancers diagnosed from April 2014 onwards (n = 28,146; sensitivity analysis 1), 45.6% (n = 12,835 patients) were found to have utilised trastuzumab. Reduced utilisation with increasing deprivation was still present (IMD 5 vs IMD 1; mvOR 0.95, 95% CI 0.86, 1.04), though this association was no longer significant (LRT p = 0.136) and was of a smaller magnitude (Appendix 3.8). When restricting the cohort to breast cancers with a HER2 status positive classification only (n = 27,712; sensitivity analysis 2) a stronger deprivation effect was observed in trastuzumab utilisation to the main model (IMD 5 vs IMD 1; mvOR 0.89, 95% CI 0.80, 0.98) but again, overall, this was not significant (LRT p = 0.073) (Appendix 3.8). When restricting the cohort to women aged ≥60 (n = 20,631; sensitivity analysis 3), a larger magnitude of reduced utilisation of trastuzumab with increasing deprivation to the main model was observed (IMD 5 vs IMD 1; mvOR 0.76, 95% CI 0.68, 0.86) and this was also significant (LRT p ≤ 0.05); in this age group 6,610 women (32.0%) utilised trastuzumab (Appendix 3.8). In contrast, when restricting the cohort to women aged <60 (n = 19,548; sensitivity analysis 3), those women residing in the most deprived areas, were more likely to utilise trastuzumab (IMD 5 vs IMD 1; mvOR 1.06, 95% CI 0.96, 1.18) though this was not significant (LRT p = 0.60) (Appendix 3.8). Overall trastuzumab utilisation was 56.6%. Finally, in the last sensitivity analysis exploring a restricted cohort with a SACT record in the time range (n = 21,881; sensitivity analysis 4) a similar result to the primary analysis of all women with a HER2+ breast cancer was observed – that is, a reduced utilisation with increasing deprivation (IMD 5 vs IMD 1; mvOR 0.87, 95% CI 0.78, 0.98). The magnitude of this effect was larger (women in the most deprived areas were 13% less likely to utilise trastuzumab) and this result was statistically significant (LRT p ≤ 0.05) (Appendix 3.9).

Multivariable Model Other Variables & their Associations with Utilisation: Several other variables in the multivariable model, other than deprivation, had statistically significant associations with reduced trastuzumab utilisation. These were: an older age (70-79 vs 80+; mvOR 7.02, 95% CI 6.12, 8.04); a low-grade tumour classification (low vs poor differentiated grade; mvOR 0.12, 95% CI 0.11, 0.14); and three or more comorbidities (3+ vs 0 comorbidities; mvOR 0.40, 95% CI 0.34, 0.47). Over the study time period, trastuzumab utilisation generally increased (2012 vs 2017; mvOR 0.41, 95% CI 0.38, 0.45). There were no clear rural/urban effects though regional variations were observed.

Table 3.3 Likelihood (OR with 95% CI and p values from logistic regression) of utilising trastuzumab by deprivation and adjusted for: age, diagnosis year, ethnicity, rural/urban categorisation, government region, stage, grade, ER status, comorbidities, and whether discussed at MDT for women with HER2+ breast cancer diagnosed between 01/01/2012 - 31/12/2017 (n = 40,179)

				Unadjusted			**Adjusted**		
	Number (%) Utilising Trastuzumab n = 17,674 (43.99)	Number (%) not Utilising Trastuzumab n = 22,505 (56.01)	P Value[1]	OR	95% CI	P Value[2]	OR	95% CI	P Value[2]
IMD[3]			0.029			**0.029**			**0.040**
1 (Least Deprived)	4,088 (44.09)	5,184 (55.91)		1.00	----- -----	-------	1.00	----- -----	-------
2	3,959 (44.15)	5,008 (55.85)		1.00	0.95 – 1.06	0.934	0.99	0.93 – 1.06	0.846
3	3,456 (42.52)	4,671 (57.48)		0.94	0.88 – 1.00	0.038	0.92	0.86 – 0.99	0.021
4	3,266 (44.28)	4,109 (55.72)		1.01	0.95 – 1.07	0.801	0.93	0.86 – 1.00	0.040
5 (Most Deprived)	2,905 (45.12)	3,533 (54.88)		1.04	0.98 – 1.11	0.200	0.92	0.85 – 0.99	0.036
Age (Years)			<0.001			**<0.001**			**<0.001**
<40	1,707 (65.25)	909 (34.75)		29.05	25.11 – 33.61	<0.001	30.55	26.22 – 35.59	<0.001
40–49	4,158 (58.12)	2,996 (41.88)		21.47	18.85 – 24.46	<0.001	24.89	21.72 – 28.52	<0.001
50–59	5,199 (53.17)	4,579 (46.83)		17.56	15.46 – 19.96	<0.001	21.21	18.56 – 24.23	<0.001
60–69	4,372 (45.57)	5,222 (54.43)		12.95	11.40 – 14.72	<0.001	16.12	14.11 – 18.41	<0.001
70–79	1,961 (30.29)	4,514 (69.71)		6.72	5.89 – 7.67	<0.001	7.02	6.12 – 8.04	<0.001
80+	277 (6.07)	4,285 (93.93)		1.00	----- -----	-------	1.00	----- -----	-------
Diagnosis Year			<0.001			**<0.001**			**<0.001**
2012	1,567 (33.23)	3,149 (66.77)		0.64	0.59 – 0.69	<0.001	0.41	0.38 – 0.45	<0.001
2013	2,575 (44.82)	3,170 (55.18)		1.04	0.97 – 1.11	0.259	0.80	0.74 – 0.86	<0.001
2014	2,866 (46.08)	3,353 (53.92)		1.09	1.02 – 1.17	0.007	0.91	0.85 – 0.99	0.020
2015	3,235 (46.35)	3,745 (53.65)		1.11	1.04 – 1.18	0.002	1.01	0.93 – 1.08	0.884
2016	3,651 (46.20)	4,251 (53.80)		1.10	1.03 – 1.17	0.003	1.07	0.99 – 1.14	0.076
2017	3,780 (43.87)	4,837 (56.13)		1.00	----- -----	-------	1.00	----- -----	-------

Table 3.3 Continued

				Unadjusted			Adjusted		
	Number (%) Utilising Trastuzumab n = 17,674 (43.99)	Number (%) not Utilising Trastuzumab n = 22,505 (56.01)	P Value[1]	OR	95% CI	P Value[2]	OR	95% CI	P Value[2]
Ethnicity			<0.001			**<0.001**			**<0.001**
White	15,566 (44.13)	19,710 (55.87)		1.00	----- -----	-------	1.00	----- -----	-------
Other Ethnic Group[4]	1,602 (51.84)	1,488 (48.16)		1.36	1.27 – 1.47	<0.001	0.91	0.84 – 1.00	0.040
Missing/unknown[5]	506 (27.91)	1,307 (72.09)		0.49	0.44 – 0.54	<0.001	0.44	0.39 – 0.50	<0.001
Rural/Urban Indicator			<0.001			**<0.001**			**<0.001**
Rural Village, Hamlet & Isolated Dwellings	1,923 (43.96)	2,451 (56.04)		0.92	0.86 – 0.98	0.013	0.88	0.81 – 0.97	0.007
Rural Town & Fringe	1,856 (43.21)	2,439 (56.79)		0.89	0.83 – 0.95	0.001	0.90	0.82 – 0.98	0.017
Urban City & Town	7,787 (42.63)	10,478 (57.37)		0.87	0.83 – 0.91	<0.001	0.84	0.78 – 0.89	<0.001
Urban Conurbation	6,108 (46.12)	7,137 (53.88)		1.00	----- -----	-------	1.00	----- -----	-------
Government Region			<0.001			**<0.001**			**<0.001**
North West	2,795 (44.30)	3,514 (55.70)		1.00	----- -----	-------	1.00	----- -----	-------
North East	981 (38.74)	1,551 (61.26)		0.80	0.72 – 0.87	<0.001	0.65	0.58 – 0.73	<0.001
West Midlands	1,725 (40.10)	2,577 (59.90)		0.84	0.78 – 0.91	<0.001	0.72	0.66 – 0.79	<0.001
Yorkshire & the Humber	1,943 (54.58)	1,617 (45.42)		1.51	1.39 – 1.64	<0.001	1.28	1.17 – 1.41	<0.001
East Midlands	1,699 (46.48)	1,956 (53.52)		1.09	1.01 – 1.19	0.035	1.07	0.97 – 1.18	0.150
East of England	2,161 (41.65)	3,027 (58.35)		0.90	0.83 – 0.97	0.004	0.79	0.72 – 0.86	<0.001
South East	2,773 (42.77)	3,710 (57.23)		0.94	0.88 – 1.01	0.081	0.90	0.83 – 0.98	0.020
South West	1,881 (43.25)	2,468 (56.75)		0.96	0.89 – 1.04	0.283	0.98	0.89 – 1.08	0.629
London	1,716 (45.15)	2,085 (54.85)		1.03	0.95 – 1.12	0.408	0.83	0.75 – 0.92	<0.001

Table 3.3 Continued

				Unadjusted			Adjusted		
	Number (%) Utilising Trastuzumab n = 17,674 (43.99)	Number (%) not Utilising Trastuzumab n = 22,505 (56.01)	P Value[1]	OR	95% CI	P Value[2]	OR	95% CI	P Value[2]
Stage			<0.001			**<0.001**			**<0.001**
I	4,680 (35.74)	8,414 (64.26)		1.00	----- -----	-------	1.00	----- -----	-------
II	8,189 (49.14)	8,474 (50.86)		1.74	1.66 – 1.82	<0.001	1.71	1.62 – 1.80	<0.001
III	2,698 (55.53)	2,161 (44.47)		2.24	2.10 – 2.40	<0.001	2.23	2.07 – 2.41	<0.001
IV	1,062 (44.83)	1,307 (55.17)		1.46	1.34 – 1.60	<0.001	1.74	1.57 – 1.92	<0.001
Unknown[6]	1,045 (32.72)	2,149 (67.28)		0.87	0.81 – 0.95	0.001	1.15	1.05 – 1.27	0.004
Grade			<0.001			**<0.001**			**<0.001**
Well Differentiated (Low Grade)	251 (11.53)	1,926 (88.47)		0.11	0.10 – 0.12	<0.001	0.12	0.11 – 0.14	<0.001
Moderately Differentiated	6,302 (36.24)	11,089 (63.76)		0.47	0.46 – 0.49	<0.001	0.53	0.50 – 0.55	<0.001
Poorly Differentiated	10,770 (54.49)	8,994 (45.51)		1.00	----- -----	-------	1.00	----- -----	-------
Other[7]	351 (41.44)	496 (58.56)		0.59	0.51 – 0.68	<0.001	0.66	0.56 – 0.77	<0.001
Multiple Tumours			<0.001			**<0.001**			-------
1	15,999 (45.32)	19,304 (54.68)		1.00	----- -----	-------	-------	----- -----	-------
>1	1,675 (34.35)	3,201 (65.65)		0.63	0.59 – 0.67	<0.001	-------	----- -----	-------
ER Status			<0.001			**<0.001**			**<0.001**
Positive	9,878 (38.78)	15,593 (61.22)		1.00	----- -----	-------	1.00	----- -----	-------
Negative	4,855 (56.45)	3,746 (43.55)		2.05	1.95 – 2.15	<0.001	1.81	1.71 – 1.92	<0.001
Unknown	2,941 (48.16)	3,166 (51.84)		1.47	1.39 – 1.55	<0.001	1.36	1.27 – 1.45	<0.001

Table 3.3 Continued

				Unadjusted			Adjusted		
	Number (%) Utilising Trastuzumab n = 17,674 (43.99)	Number (%) not Utilising Trastuzumab n = 22,505 (56.01)	P Value[1]	OR	95% CI	P Value[2]	OR	95% CI	P Value[2]
CCM (Between 78 to 6 Months Prior to Diagnosis)			<0.001			**<0.001**			**<0.001**
0	15,414 (46.85)	17,485 (53.15)		1.00	----- -----	-------	1.00	----- -----	-------
1-2	2,044 (34.63)	3,858 (65.37)		0.60	0.57 – 0.64	<0.001	0.84	0.78 – 0.90	<0.001
3+	216 (15.67)	1,162 (84.33)		0.21	0.18 – 0.24	<0.001	0.40	0.34 – 0.47	<0.001
Discussed at MDT			<0.001			**<0.001**			**<0.001**
Yes	12,936 (46.56)	14,847 (53.44)		1.00	----- -----	-------	1.00	----- -----	-------
No	2,536 (40.66)	3,701 (59.34)		0.79	0.74 – 0.83	<0.001	0.78	0.73 – 0.84	<0.001
Missing	2,202 (35.75)	3,957 (64.25)		0.64	0.60 – 0.68	<0.001	0.52	0.49 – 0.56	<0.001

[1]Chi-square P value
[2]Bolded P Values are from LRT of the variable's contribution to the model. Unbolded P values are from a test of whether the OR is different from 1.
[3]For diagnosis year 2012, IMD_2010 was used and for diagnosis years 2013 - 2017, IMD_2015 was used.
[4]Other refers to Asian/British Asian, Black/African/Caribbean/Black British, mixed/multiple ethnic groups and other ethnic groups.
[5]Missing ethnicity refers to missing and unknown ethnicity classifications.
[6]Unknown staging refers to missing and unstageable tumours.
[7]Other refers to undifferentiated or anaplastic, undetermined, and missing tumour grades. Abbreviations: CCM: Charlson Comorbidity Index; ER: oestrogen receptor status; IMD: Index of Multiple Deprivation; LRT: likelihood ratio test; n: Number; MDT: multidisciplinary team; OR: odds ratio; 95% CI: 95% confidence interval.

3.6.3 Secondary Analysis: Conventional Therapy Utilisation (Surgery)

Descriptive Statistics: 75% (n = 30,124) of women had breast cancer directed surgery. Reduced utilisation of breast cancer directed surgery was associated with increasing deprivation: 76% in IMD 1 (least deprived); 78% in IMD 2; 75% in IMD 3; 73% in IMD 4; and 71% in IMD 5 (most deprived) (Appendix 3.10a). The proportion of women utilising surgery over the time frame covered by SACT slightly dropped (80% in 2012 compared to 69% in 2017), despite diagnosed cases rising over the period in question (Appendix 3.10b). Surgical use was high in stage I-III cancers and 91% of patients with stage I cancers had surgery (Appendix 3.10c).

Multivariable Model Deprivation Associations: Utilisation of breast cancer directed surgery generally decreased as the level of deprivation increased from 76.3% in (IMD 1 least deprived), to 71.4% (IMD 5 most deprived). This trend for decreasing surgical utilisation with increasing deprivation level persisted in the final multivariable model where a strong and significant (LRT $p \leq 0.05$) association was observed (Figure 3.8 and Table 3.4). There was a statistically significant linear trend ($p \leq 0.05$) across all deprivation categories. Though the most pronounced effect were seen in IMD 5 (most deprived women) (IMD 5 vs IMD 1; mvOR 0.79, 95% CI 0.73, 0.86). This meant that women living in the most deprived areas were 21% less likely to utilise breast cancer directed surgery compared to those residing in the least deprived areas.

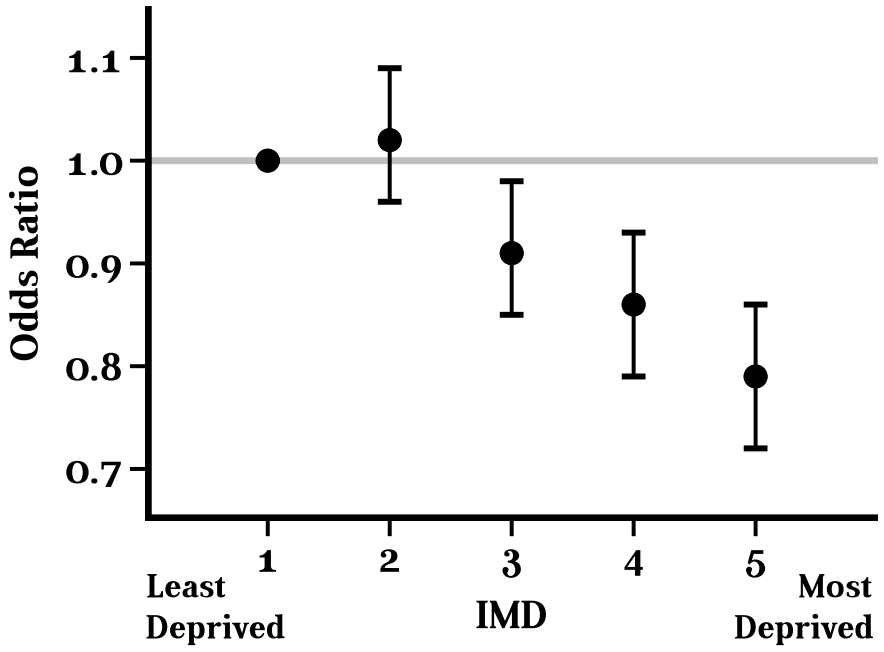


Figure 3.8 Breast cancer directed surgery utilisation by deprivation (IMD 1 least deprived; IMD 5 most deprived). Surgery multivariable model adjusted for: age and stage, diagnosis year, ethnicity, rural/urban categorisation, grade, ER status, comorbidities, and whether discussed at MDT. Abbreviations: IMD: Index of multiple deprivation; ER: Oestrogen receptor; MDT: Multidisciplinary team.

Sensitivity Analyses: When restricting the cohort to breast cancers diagnosed from April 2014 onwards (n = 28,146; sensitivity analysis 1), 73.2% (n = 20,594 women) were found to have utilised breast cancer directed surgery (Appendix 3.11). Again, a significant association (LRT $p \leq 0.05$), of near similar magnitude to the main model was found. This showed reduced surgery utilisation as deprivation level increased (IMD 5 vs IMD 1; mvOR 0.82, 95% CI 0.74, 0.90). Restricting the cohort to a HER2 status positive classification only (n = 27,712; sensitivity analysis 2) found that 75% of patients utilised surgery (n = 20,666) and that the same patterns of reduced surgery utilisation as deprivation status increased as seen in the main model were replicated (IMD 5 vs IMD 1; mvOR 0.80, 95% CI 0.72, 0.88) (Appendix 3.11). Again, these associations were significant (LRT $p \leq 0.05$) and of comparable magnitude to the main model. Finally, with restriction of the cohort by age (sensitivity analysis 3), it was observed that in women <60 years old (n = 19,548), 76% (n = 14,922) utilised surgery. There was also reduced utilisation with increasing deprivation status (IMD 5 vs IMD 1; mvOR 0.88, 95%] CI 0.78, 1.00) (Appendix 3.11). This association remained significant (LRT $p \leq 0.05$). Finally, in women aged ≥60 years (n = 20,631), an even strong association of reduced surgery utilisation with increasing deprivation was observed than the main model (IMD 5 vs IMD 1; mvOR 0.71, 95% CI 0.63, 0.80); 73.7% of women in this age group utilised surgery (Appendix 3.11) and this was significant (LRT $p \leq 0.05$).

Multivariable Model Other Variables & their Associations with Utilisation: A few other factors, other than deprivation, were also associated with statistically significant poorer surgical utilisation. These included: a non-white ethnicity (other ethnic group vs white; mvOR 0.75, 95% CI 0.68, 0.82); and three or more comorbidities (3+ vs 0 comorbidities; mvOR 0.47, 95% CI 0.41, 0.53). Generally, early-stage patients observed increasing utilisation with increasing age (age 60-69 early stage + vs 70+ early stage; mvOR 2.23, 95% CI 2.05, 2.43), whereas patient with metastatic or unknown staging showed decreasing utilisation trends with increasing age (age 70+ metastatic or unknown stage vs 70+ early stage; mvOR 0.08 [95%], 0.07, 0.09). Surgical utilisation decreased over the period in question despite an increase in diagnoses (2012 vs 2017; mvOR 2.72 [95%], 2.46, 3.01) (Table 3.4).

Table 3.4 Likelihood (OR with 95% CI and p values from logistic regression) of utilising breast cancer directed surgery by deprivation and adjusted for: age and stage, diagnosis year, ethnicity, rural/urban categorisation, grade, ER status, comorbidities, and whether discussed at MDT for women with HER2+ breast cancer diagnosed between 01/01/2012 – 31/12/2017 (n = 40,179)

				Unadjusted			**Adjusted**		
	Number (%) Utilising Surgery n = 30,124 (74.97)	Number (%) not Utilising Surgery n = 10,055 (25.03)	P Value[1]	OR	95% CI	P Value[2]	OR	95% CI	P Value[2]
IMD[3]			<0.001			**<0.001**			**<0.001**
1 (Least Deprived)	7,075 (76.31)	2,197 (23.69)		1.00	----- -----	-------	1.00	----- -----	-------
2	6,948 (77.48)	2,019 (22.52)		1.07	1.00 – 1.14	0.059	1.02	0.94 – 1.10	0.618
3	6,109 (75.17)	2,018 (24.83)		0.94	0.88 – 1.01	0.081	0.91	0.84 – 0.98	0.016
4	5,397 (73.18)	1,978 (26.82)		0.85	0.79 – 0.91	<0.001	0.86	0.79 – 0.93	<0.001
5 (Most Deprived)	4,595 (71.37)	1,843 (28.63)		0.77	0.72 – 0.83	<0.001	0.79	0.73 – 0.86	<0.001
Age (Years) & Stage[4]			<0.001			**<0.001**			**<0.001**
Age <50 Stage Early	6,639 (78.42)	1,827 (21.58)		1.16	1.08 – 1.24	<0.001	1.12	1.04 – 1.21	0.004
Age <50 Stage Metastatic or Unknown	516 (39.57)	788 (60.43)		0.21	0.18 – 0.24	<0.001	0.20	0.18 – 0.23	<0.001
Age 50 – 59 Stage Early	7,305 (84.77)	1,312 (15.23)		1.77	1.64 – 1.91	<0.001	1.74	1.60 – 1.88	<0.001
Age 50 – 59 Stage Metastatic or Unknown	462 (39.79)	699 (60.21)		0.21	0.19 – 0.24	<0.001	0.20	0.17 – 0.23	<0.001
Age 60 – 69 Stage Early	7,566 (88.06)	1,026 (11.94)		2.34	2.16 – 2.54	<0.001	2.23	2.05 – 2.43	<0.001
Age 60 – 69 Stage Metastatic or Unknown	390 (38.92)	612 (61.08)		0.20	0.18 – 0.23	<0.001	0.18	0.16 – 0.21	<0.001
Age 70+ Stage Early	6,784 (75.88)	2,157 (24.12)		1.00	----- -----	-------	1.00	----- -----	-------
Age 70+ Stage Metastatic or Unknown	462 (22.04)	1,634 (77.96)		0.09	0.08 – 0.10	<0.001	0.08	0.07 – 0.09	<0.001
Diagnosis Year			<0.001			**<0.001**			**<0.001**
2012	3,793 (80.43)	923 (19.57)		1.88	1.73 – 2.05	<0.001	2.72	2.46 – 3.01	<0.001
2013	4,515 (78.59)	1,230 (21.41)		1.68	1.56 – 1.82	<0.001	2.28	2.08 – 2.50	<0.001
2014	4,811 (77.36)	1,408 (22.64)		1.57	1.45 – 1.69	<0.001	1.80	1.66 – 1.97	<0.001
2015	5,259 (75.34)	1,721 (24.66)		1.40	1.30 – 1.50	<0.001	1.45	1.34 – 1.57	<0.001
2016	5,836 (73.85)	2,066 (26.15)		1.29	1.21 – 1.38	<0.001	1.35	1.25 – 1.46	<0.001
2017	5,910 (68.59)	2,707 (31.41)		1.00	----- -----	-------	1.00	----- -----	-------

Table 3.4 Continued

	Number (%) Utilising Surgery n = 30,124 (74.97)	Number (%) not Utilising Surgery n = 10,055 (25.03)	P Value[1]	Unadjusted OR	Unadjusted 95% CI	Unadjusted P Value[2]	Adjusted OR	Adjusted 95% CI	Adjusted P Value[2]
Ethnicity			<0.001			**<0.001**			**<0.001**
White	26,804 (75.98)	8,472 (24.02)		1.00	----- -----	-------	1.00	----- -----	-------
Other Ethnic Group[5]	2,107 (68.19)	983 (31.81)		0.68	0.63 – 0.73	<0.001	0.75	0.68 – 0.82	<0.001
Missing/unknown[6]	1,213 (66.91)	600 (33.09)		0.64	0.58 – 0.71	<0.001	0.77	0.69 – 0.87	<0.001
Rural/Urban Indicator			<0.001			**<0.001**			**<0.001**
Rural Village, Hamlet & Isolated Dwellings	3,443 (78.72)	931 (21.28)		1.46	1.35 – 1.58	<0.001	1.31	1.19 – 1.44	<0.001
Rural Town & Fringe	3,288 (76.55)	1,007 (23.45)		1.29	1.19 – 1.40	<0.001	1.16	1.06 – 1.28	0.002
Urban City & Town	13,898 (76.09)	4,367 (23.91)		1.26	1.19 – 1.32	<0.001	1.20	1.14 – 1.28	<0.001
Urban Conurbation	9,495 (71.69)	3,750 (28.31)		1.00	----- -----	-------	1.00	----- -----	-------
Government Region			<0.001			**<0.001**			-------
North West	4,889 (77.49)	1,420 (22.51)		1.00		-------	-----	----- -----	-------
North East	1,968 (77.73)	564 (22.27)		1.01	0.91 – 1.13	0.813	-----	----- -----	-------
West Midlands	3,223 (74.92)	1,079 (25.08)		0.87	0.79 – 0.95	0.002	-----	----- -----	-------
Yorkshire & the Humber	2,728 (76.63)	832 (23.37)		0.95	0.86 – 1.05	0.326	-----	----- -----	-------
East Midlands	2,735 (74.83)	920 (25.17)		0.86	0.79 – 0.95	0.003	-----	----- -----	-------
East of England	3,952 (76.18)	1,236 (23.82)		0.93	0.85 – 1.01	0.096	-----	----- -----	-------
South East	4,817 (74.30)	1,666 (25.70)		0.84	0.77 – 0.91	<0.001	-----	----- -----	-------
South West	3,422 (78.68)	927 (21.32)		1.07	0.98 – 1.18	0.144	-----	----- -----	-------
London	2,390 (62.88)	1,411 (37.12)		0.49	0.45 – 0.54	<0.001	-----	----- -----	-------
Grade			<0.001			**<0.001**			**<0.001**
Well Differentiated (Low Grade)	1,795 (82.45)	382 (17.55)		1.46	1.30 – 1.64	<0.001	1.35	1.18 – 1.53	<0.001
Moderately Differentiated	12,872 (74.02)	4,519 (25.98)		0.89	0.85 – 0.93	<0.001	0.87	0.83 – 0.92	<0.001
Poorly Differentiated	15,071 (76.25)	4,693 (23.75)		1.00	----- -----	-------	1.00	----- -----	-------
Other[7]	386 (45.57)	461 (54.43)		0.26	0.23 – 0.30	<0.001	0.35	0.30 – 0.42	<0.001

Table 3.4 Continued

				Unadjusted			Adjusted		
	Number (%) Utilising Surgery n = 30,124 (74.97)	Number (%) not Utilising Surgery n = 10,055 (25.03)	P Value[1]	OR	95% CI	P Value[2]	OR	95% CI	P Value[2]
Multiple Tumours			<0.001			**<0.001**			--------
1	26,592 (75.33)	8,711 (24.67)		1.00	----- -----	-------	-----	----- -----	-------
>1	3,532 (72.44)	1,344 (27.56)		0.86	0.80 – 0.92	<0.001	-----	----- -----	-------
ER Status			<0.001			**<0.001**			**<0.001**
Positive	19,550 (76.75)	5,921 (23.25)		1.00	----- -----	-------	1.00	----- -----	-------
Negative	6,303 (73.28)	2,298 (26.72)		0.83	0.79 – 0.88	<0.001	0.87	0.81 – 0.93	<0.001
Unknown	4,271 (69.94)	1,836 (30.06)		0.70	0.66 – 0.75	<0.001	0.89	0.83 – 0.96	0.002
CCM (Between 78 to 6 Months Prior to Diagnosis)			<0.001			**<0.001**			**<0.001**
0	25,110 (76.32)	7,789 (23.68)		1.00	----- -----	-------	1.00	----- -----	-------
1-2	4,257 (72.13)	1,645 (27.87)		0.80	0.75 – 0.85	<0.001	0.89	0.83 – 0.96	0.002
3+	757 (54.93)	621 (45.07)		0.38	0.34 – 0.42	<0.001	0.47	0.41 – 0.53	<0.001
Discussed at MDT			<0.001			**<0.001**			**<0.001**
Yes	22,164 (79.78)	5,619 (20.22)		1.00	----- -----	-------	1.00	----- -----	-------
No	4,457 (71.46)	1,780 (28.54)		0.63	0.60 – 0.68	<0.001	0.79	0.74 – 0.85	<0.001
Missing	3,503 (56.88)	2,656 (43.12)		0.33	0.32 – 0.35	<0.001	0.41	0.39 – 0.44	<0.001

[1]Chi-square P value
[2]Bolded P Values are from LRT of the variable's contribution to the model. Unbolded P values are from a test of whether the OR is different from 1.
[3]For diagnosis year 2012, IMD_2010 was used and for diagnosis years 2013 - 2017, IMD_2015 was used.
[4]Unknown staging refers to missing and unstageable tumours.
[5]Other refers to Asian/British Asian, Black/African/Caribbean/Black British, mixed/multiple ethnic groups and other ethnic groups.
[6]Missing/unknown ethnicity refers to missing and unknown ethnicity classifications.
[7]Other refers to undifferentiated or anaplastic, undetermined, and missing tumour grades. Abbreviations: CCM: Charlson Comorbidity Index; ER: oestrogen receptor status; IMD: Index of Multiple Deprivation; LRT: likelihood ratio test; MDT: multidisciplinary team; n: number; OR: odds ratio; 95% CI: 95% confidence interval.

3.6.4 Secondary Analysis: Conventional Therapy Utilisation (Chemotherapy)

Descriptive Statistics: 60.8% (n = 24,444) women utilised a chemotherapy. Little difference in utilisation of chemotherapy by deprivation quintile was observed: 59% for IMD 1 (least deprived); 60% for IMD 2; 61% for IMD 3; 61% for IMD 4; and 64% for IMD 5 (most deprived) (Appendix 3.12a). Whilst the number of breast cancer cases diagnosed increased yearly over the time frame of interest, the proportion of women utilising a chemotherapy declined from 65% of cases diagnosed in 2012 to 57% of cases diagnosed in 2017 (Appendix 3.12b). Most women with a breast cancer staged above I were treated with chemotherapy: 68% for stage II, 78% for stage III and 65% for stage IV (Appendix 3.12c).

Multivariable Model Deprivation Associations: The chemotherapy models were initially stratified by receipt of surgery and then were further described by age (≥60 and <60).

Women who had Surgery: In women receiving surgery who were aged ≥60, there was a slight decrease in utilisation with increasing deprivation status: 48.1% (IMD 1 least deprived); 48.1% (IMD 2); 47.9% (IMD 3); 46.9% (IMD 4); and 46.6% (IMD 5 most deprived). This trend for reduced chemotherapy utilisation with increasing deprivation status persisted after adjustment for confounders, though this was not significant (LRT p value = 0.72). Women residing in the most deprived areas were 5% less likely to utilise chemotherapy (IMD 5 vs IMD 1; mvOR 0.95, 95% CI 0.85, 1.07) (Appendix 3.13). The test for linear trend was not significant (p = 0.32).

In women aged <60, chemotherapy utilisation was observed to increase with increasing deprivation: 73.8% (IMD 1 least deprived); 76.6% (IMD 2); 77.3% (IMD 3); 78.2% (IMD 4); and 81.4% (IMD 5 most deprived). This trend for increasing chemotherapy utilisation with increasing deprivation persisted after controlling for confounders and was found to be significant (LRT $p \leq 0.05$). There was a slight trend for increasing utilisation with deprivation category and the test for linear trend was found to be significant ($p \leq 0.05$). The magnitude of this association was most stark when comparing those women residing in the least to most deprived areas (IMD 5 vs IMD 1; mvOR 1.36, 95% CI 1.19, 1.57) (Appendix 3.14).

Utilisation in women received surgery by age is depicted in Figure 3.9a.

Women who did not have Surgery: In women not receiving surgery who were aged ≥60, chemotherapy utilisation decreased with increasing deprivation status: 39.0% (IMD 1 least deprived); 35.5% (IMD 2); 33.3% (IMD 3); 32.4% (IMD 4); and 35.5% (IMD 5 most deprived). This association was significant (LRT $p \leq 0.05$). The trend in reduced utilisation occurred across all deprivation categories and the test for linear trend was significant ($p \leq 0.05$). The largest magnitude of difference in utilisation was observed between women residing in the least and most deprived areas (IMD 5 vs IMD 1; mvOR 0.85, 95% CI 0.70, 1.02). Women with the highest level of deprivation were found to be 15% less likely to utilise chemotherapy (Appendix 3.15).

In women aged <60 and not receiving surgery the opposite was observed: chemotherapy utilisation increased with increasing deprivation status: 75.6% (IMD 1); 77.9% (IMD 2); 86.1% (IMD 3); 83.3% (IMD 4); and 86.2% (IMD 5). This association was significant (LRT p value ≤ 0.05) and this pattern persisted generally across each deprivation category and despite the wide confidence intervals. The test for linear test for trend was also significant (p value ≤ 0.05). The most marked increase in chemotherapy utilisation was observed between women residing in the least and most deprived locations (IMD 5 vs IMD 1; mvOR 1.60, 95% CI 1.21, 2.11) (Appendix 3.16).

Utilisation of chemotherapy in women not receiving surgery by age is shown in Figure 3.9b.

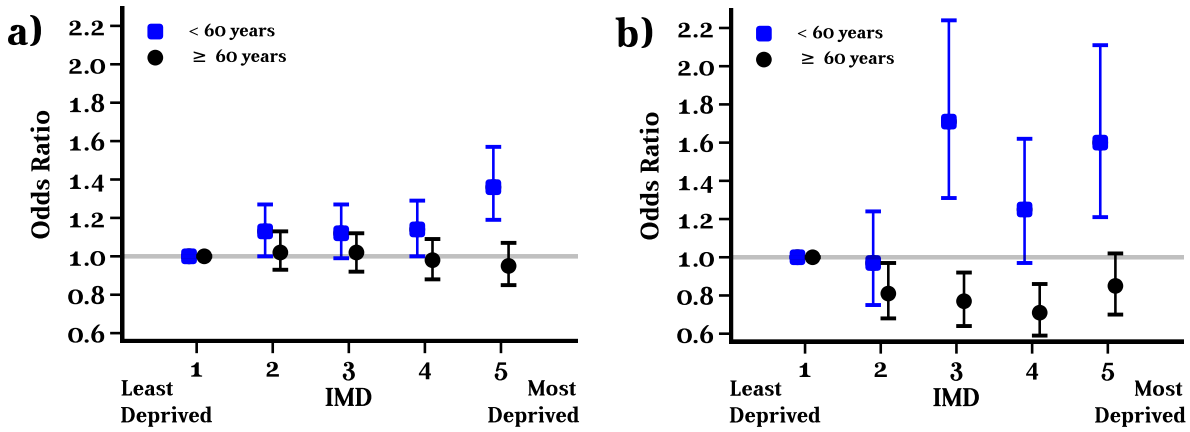


Figure 3.9 Chemotherapy utilisation in women **a)** receiving surgery and **b)** not receiving surgery by deprivation (IMD 1 least deprived; IMD 5 most deprived). All multivariable models adjusted for: diagnosis year, stage, ethnicity, comorbidities (apart from the model for women aged <60 receiving surgery), and whether discussed at MDT. Abbreviations: IMD: Index of Multiple Deprivation.; MDT: Multi-disciplinary team.

Sensitivity Analyses: As per the main chemotherapy analysis, sensitivity analyses were initially stratified by receipt of surgery and then were further described by age (≥60 and <60).

Women who had Surgery: Sensitivity analysis 1 considered restricting the cohort to breast cancers diagnosed from April 2014 onwards. For women who had surgery and aged ≥60, 46% (n = 4,873) were found to have utilised chemotherapy and the association with deprivation deviated in magnitude very little from the main model (IMD 5 vs IMD 1; mvOR 0.95, 95% CI 0.83, 1.09). Though this association was not found to be statistically significant (LRT p = 0.844) (Appendix 3.17). In women receiving surgery who received surgery and were aged <60, 74% (n =7,441) utilised chemotherapy and again associations with deprivation were like those reported in the main multivariable model. This analysis showed an increasing utilisation with increasing deprivation (IMD 5 vs IMD 1; mvOR 1.35, 95% CI 1.14, 1.59) (Appendix 3.18).

Sensitivity analysis 2 restricted the cohort to women with a HER2 status positive classification only. In women who had surgery who were aged ≥60, 59% (n =5,828 women) were found to utilise chemotherapy. The magnitude of the association of chemotherapy utilisation with deprivation was of a similar magnitude to the main model and this association was found to also not be significant (LRT p = 0.773). Overall women in the most deprived category were 3% less likely to utilise chemotherapy compared to those in the least deprived category (IMD 5 vs IMD 1; mvOR 0.97, 95% CI 0.84, 1.11) (Appendix 3.17). In women receiving surgery, aged <60, 87% were found to utilise chemotherapy (n = 9,397). As per the main model, there was an association of increasing chemotherapy utilisation with increasing deprivation status. This association was larger in magnitude than in the main model. This meant that women residing in the most deprived areas were more likely than those in the least deprived ones to utilise chemotherapy (IMD 5 vs IMD 1; mvOR 1.62, 95% CI 1.32, 1.99) and this association was significant (LRT p ≤ 0.05) (Appendix 3.18).

Women who did not have Surgery: Sensitivity analysis 1 considered restricting the cohort to breast cancers diagnosed from April 2014 onwards. For women who did not have surgery and who were aged ≥60, n =1,448 (35%) utilised chemotherapy. The pattern and magnitude of ORs followed a similar pattern to those of main model – decreasing utilisation by increasing deprivation status. This association was significant (LRT ≤ 0.05) and most contrasting when considering those women residing in the least versus most deprived areas (IMD 5 vs IMD 1; mvOR 0.85, 95% CI 0.68, 1.06) (Appendix 3.19). For women aged <60 who did not receive surgery, n = 2,849 (82%) utilised chemotherapy. Again, the pattern of increasing utilisation

with increasing deprivation with similar ORs to the main model was observed and found to be significant (LRT ≤ 0.05). This meant that women residing in the most deprived areas were more likely to utilise chemotherapy those than women in the least deprived areas (IMD 5 vs IMD 1; mvOR 1.62, 95% CI 1.16, 2.26) (Appendix 3.20).

Sensitivity analysis 2 restricted the cohort to women with a HER2 status positive classification only. For women who did not have surgery and who were aged ≥60, n = 1,544 (46%) utilised chemotherapy. Once again, decreased utilisation was associated with increasing deprivation. Whilst the magnitude of the association was like the main model, the association was not found to be significant (LRT p = 0.069). Associations were largest when comparing the most to least deprived women (IMD 5 vs IMD 1; mvOR 0.86, 95% CI 0.68, 1.09) (Appendix 3.19). Finally, in women who did not have surgery and were aged <60, n = 3,142 (11%) utilised chemotherapy. These associations between deprivation status and chemotherapy utilisation were found to be of a much greater magnitude to the main model and they were significant (LRT p ≤ 0.05). Apart from IMD 3, there was an increasing utilisation of chemotherapy with increasing deprivation status (IMD 5 vs IMD 1; mvOR 2.08, 95% CI 1.48, 2.92) (Appendix 3.20).

3.7 Discussion

3.7.1 Main Findings

This population-based cohort study (n = 40,179 women) investigated associations between deprivation and breast cancer treatment utilisation in women with HER2+ breast cancer diagnosed in England between 01/01/2012 - 31/12/2017. Residence in an area of high deprivation was associated with a reduced likelihood of treatment utilisation. For trastuzumab utilisation, modest inequalities were observed; women residing in areas of higher deprivation were 8% less likely to utilise trastuzumab compared to those residing in the least deprived areas. This association was significant and from IMD 3 onwards was found to be marked with reduced utilisation. Inequalities in breast cancer directed surgery utilisation were more pronounced than observed with trastuzumab utilisation as 21% of patients residing in areas of high deprivation were less likely to utilise breast cancer directed surgery compared with those women living in affluent areas. This association was significant and occurred between each increasing deprivation category. Chemotherapy utilisation was more complex and varied by whether the patient had also received surgery - as well as their age. Women aged ≥60 (regardless of whether they had received surgery or not) and who resided in areas of higher level of deprivation experienced a reduced likelihood of chemotherapy utilisation. This was not the case for women aged <60, who experienced an increased likelihood of chemotherapy utilisation (especially if

they had also received surgery) with residence in an area of higher deprivation compared to residence in the least deprived areas. All chemotherapy model utilisation associations with deprivation were found to be significant apart from the model exploring women aged ≥60 who also had surgery.

The finding that there are socio-economic inequalities in: (i) conventional breast cancer treatment; and (ii) trastuzumab is consistent with the previous literature (92,216–219) as well as the conclusions stated in Chapter 2 (Appendix 2.1) (231) - that conventional cancer treatment inequalities associated with a low SES have now infiltrated novel anti-cancer therapies. The work presented in this Chapter further strengthens the findings from the breast cancer meta-analysis (Chapter 2) (IMD 5 vs IMD 1; pooled OR = 0.93, 95% CI 0.78, 1.10), suggesting that inequalities associated with a low SES in treatment utilisation are not necessarily healthcare system specific given their endurance (at a similar magnitude) in this observational study (IMD 5 vs IMD 1; mvOR 0.92, 95% CI 0.85,0.99) where the healthcare system was publicly funded.

This analysis represents the first English description of trastuzumab utilisation using SACT data. In summary, the findings of this work suggest that SES has a persistent role in influencing treatment utilisation in HER2+ breast cancer despite the publicly funded nature of English healthcare - though the magnitude of the inequalities observed were larger with conventional breast cancer directed surgery treatment than with the novel anti-cancer therapy, trastuzumab. Given the potential of novel anti-cancer therapies, such as trastuzumab, to improve patient outcomes (e.g. survival), these findings are concerning for patients, the NHS and wider society. This is because SES should not be a factor in treatment receipt, especially in a publicly funded healthcare system. Hence efforts will be needed to address these inequalities to ensure fairer access to treatments moving forward.

3.7.2 Interpretation of Findings

The findings suggest that despite advances in the treatment of HER2+ breast cancer, treatment utilisation is still associated with patients' SES. This result fits with the literature documenting reduced utilisation of conventional breast cancer treatments in low SES patients (216–220) along with wider research detailing engrained English health inequalities in cancer (121,122). If hypotheses from several studies which suggest that SES differences in the utilisation of treatments may account for survival and mortality variation are indeed true (126,213), then the observed inequalities noted here may have important outcomes for patients both receiving and not receiving trastuzumab.

The observation that inequalities in trastuzumab utilisation with a low SES were present raises the question as to why this is the case given the promise of precision medicine and its ability to simplify clinical decision making based on predictive biomarker status (presence of HER2+ in this instance). However, as discussed in Chapter 1, inequality causation is complex and multifactorial. It therefore seems probable that novel anti-cancer therapies are only part of the approach needed to tackle the wider, engrained cancer health inequalities in England. Precision medicines, such as trastuzumab, likely overstate the importance of genetics and downplay the wider role that of the SDoH play as a barrier to fair treatment access (441). As trastuzumab prescribing is a "downstream" intervention, targeting only the final stage of treatment receipt, a focus on precision medicine development alone fails to consider the wider "fundamental causes" or "upstream" factors (unequal distribution of income, wealth, and power) that are also important in understanding why a patient may or may not receive a treatment (27). Likewise, it is possible that the patients may have declined treatment out of choice and that such decisions may be socio-economically patterned. Furthermore, drug utilisation may be complicated further by decision making in an MDT capacity - which too may be subject to biases.

As associations of a low SES with treatment utilisation varied on account of treatment type (less stark for trastuzumab than for breast cancer surgery utilisation), this study provides some evidence that even if novel anti-cancer therapies are not a solution to English cancer treatment inequalities, they may be useful for minimising its exacerbation. There may be several factors which may explain why trastuzumab utilisation had fewer stark associations with a low SES than with conventional breast cancer treatments (surgery and chemotherapy) - especially in a publicly funded English NHS, where an individual's ability to pay is not a barrier to treatment receipt (442). Firstly, clinical factors may be important. For example, as trastuzumab is generally well tolerated in most women, toxicity is often lower than with cytotoxic chemotherapy therapy, and monitoring is (or should be) in place to minimise potential associated cardiovascular complications (443). The proportion women therefore potentially able to tolerate trastuzumab is wide. Second, healthcare system factors may be key. This is because testing for HER2+ receptor status, as outlined in NICE guidance (166), is routine practice - and has been for some time. Established testing means that trastuzumab clinical decision making is now guided by predictive biomarker test results (presence or absence of the HER2 receptor). This therefore reduces the probability that bias and stereotyping (for example based on characteristics such as SES) (227) will influence treatment utilisation. Decision making in this treatment context is thus very different to decision making for breast cancer directed surgery. In the latter, emphasis is placed on clinical judgment and whilst key

performance indicators (KPIs) are in place to ensure some level of consistency, surgical decision making is understandably more nuanced. This means that factors such as the Trust's surgical provision and case volume, along with the surgeon's skillsets/expertise can, and do, influence treatment utilisation (444–446). As these factors are very different to those at play than with trastuzumab decision making, this may, in part, help explain some of the observed discrepancies in the strength of the associations with SES between different treatment types (surgical uptake differences by SES were more pronounced than with trastuzumab). It also seems unlikely that stage of diagnosis explains the differences observed with trastuzumab and surgery utilisation by low SES as both multivariate models were adjusted for stage. Despite this, there were differences in the associations between stage and treatment between the models which may be important. In the surgery model, earlier stage was associated with significantly increased utilisation of surgery when compared to metastatic disease (much less likely). In contrast, the likelihood of trastuzumab utilisation was high in all stages other than stage I. Stage therefore appears more a consideration to breast cancer directed surgery decision-making than with trastuzumab utilisation.

The less pronounced inequalities observed with trastuzumab, when compared to surgery utilisation also questions the importance of time in facilitating intervention success. This may be because the introduction of a new intervention risks the possibility of benefiting those of high SES with resources to gain priority access (a so called IGI - see Chapter 1) (56). This risk of variation between those who have and those who do not have resources for access is particularly a concern when treatments are new (see the Inverse Inequity Hypothesis) (230). It was not possible in this analysis to assess whether, when trastuzumab was introduced into clinical guidelines, utilisation was less in patients with a low SES who had fewer resources to gain priority access. However, as trastuzumab is now an established treatment, the role of time, may in part, explain the minimal inequalities observed here as overtime, inequalities in utilisation are less (although they persist). Or as Victori *et al* (2000), hypothesise, that the medical intervention has had time to "becomes standard practice" so that all can benefit equally from the drug's introduction (230). This may mean that in cancers, such as lung, where treatments are continually evolving, that precision medicines may provide more of a barrier to fair utilisation than those medicines with established practice and clear NICE guidance. Well established practices may also extend to engrained predictive biomarker testing. For example, in breast cancer, it is not just HER2+ screening that is routine and well established - but also ER and PR testing. If appropriate identification of a targetable tumour is the first step towards drug utilisation, a lack of an established testing screen (e.g. in lung cancer where new

biomarkers are continually being discovered) may serve to increase inequalities. This hurdle of obtaining a first predictive biomarker test as a barrier to treatment utilisation is discussed again further in Chapters 4 and 5. Additionally, it is possible that women not utilising trastuzumab did in fact utilise another novel anti-cancer therapy (e.g. pertuzumab) which may also be important for consideration in discussion around the extent to which IGI exist.

There is no definitive or clear explanation for the apparently counterintuitive finding that increasing deprivation is associated with increased chemotherapy utilisation in younger women (both with and without surgery). The findings potentially indicate that treatment ordering, in addition to other known factors such as stage, may be important in determining treatment use. For most patients with HER2+ breast cancer, surgery, where possible, is first line treatment - and, as shown here, surgical receipt is not equitable. Applying a surgical "filter" may alter the demographics (leaving a larger pool of deprived, younger women fit for chemotherapy utilisation), which may, in part, explain the increased chemotherapy use observed. Although additional factors such as a patient's willingness to have chemotherapy when advised by an oncologist along with presentation of a more advanced cancer (still within the same stage) may also be important for explaining such associations too. It is also possible that as younger women are more likely to develop more aggressive breast cancer sub-types, including HER2+ tumours associated with poor prognosis (447), that this may also have influenced the treatment patterns seen here. If this is the case, and treatment ordering is important, future studies may also need to consider this to avoid obscuring true inequalities in treatment utilisation. More patients receiving trastuzumab also received surgery than did not, hence addressing barriers to fair trastuzumab utilisation may also in part start with fair access to surgery. Alternatively, it is entirely plausible that interventions/policies targeting health inequality reduction with chemotherapy utilisation are already having beneficial effects in select patient groups, hence the increased utilisation reported here. Further investigation of this result is required.

There were other interesting findings in these results. For example, the observation that HER2+ breast cancer diagnoses are rising over time. This has been found in other breast cancer studies too (448). The first possibility is to consider whether this could be ascertainment, however this seems unlikely as there is no evidence that ascertainment of breast cancer in general by the NCRD has changed over time. Rather, it has long been very high as is recognised by its inclusion in the IARC Series Cancer Incidence in Five Continents (limited to registries which meet international quality standards) (449). Increased breast cancer screening programme uptake could therefore provide a possible explanation for this observation of rising diagnoses

observed here (450). It was possible to explore the role of screening over time on this data as a route to diagnosis variable was obtained during the initial ODR request (this variable was not used in statistical models due to substantial missing data for the diagnosis year 2017). However, the data (Appendix 3.21) shows that proportions of screening-detected cancers fluctuated very little year on year over the time frame of the analysis and in fact accounted for 22.6% of all routes to diagnosis in 2012 as well as 2016. This finding fits with other studies suggesting that screening account for little of the long-term increase in breast cancer incidence (448). Having ruled out screening as a possible cause for the increasing diagnoses, changes in testing practice provide an alternative hypothesis. Appendix 3.22 examines HER2 status classification over time. It is striking to note that all the increase in HER2 diagnoses come from a rise in borderline HER2 classifications. Over the duration of the dataset, the number of a definitive positive HER2 out of all HER2 status classifications remains approximately constant, fluctuating from 4,247 people in 2012 to 4,493 people in 2017 with a high of 4,891 people in 2013. In contrast, the borderline HER2 statuses out of all HER2 status classifications increase year on year from 469 people in 2012 to 3,936 people in 2017. Reasoning for the increase in borderline cases is unknown but may result from changing interpretation of testing data.

Other risk factors for treatment utilisation were also identified. Of statistical significance, was the finding that non-white women were 25% less likely to utilise breast cancer directed surgery than white women. Similar trends by ethnicity were observed with trastuzumab utilisation but these were less stark (non-white patients were 9% less likely to utilise). These finding support evidence documenting that cancer healthcare experiences varies by ethnicity. For example, Black and Asian patients wait longer than white patients to receive a cancer diagnosis (451) and ethnic minority patients report overall worse cancer care quality experiences (452). It is not clear why non-white women were less likely to utilise treatments and previous work on this topic has found associations to be complex and poorly understood (451). One possible explanation is that ethnic treatment inequalities reflect broader economic and social inequalities (including socio-economic inequalities), driven by racial discrimination and engrained biases – and these mechanisms also filter into decisions regarding access to treatment (453). Though as models were adjusted for SES, socio-economic explanations do not account fully for the ethnic inequalities observed here. On the other hand, it may not be the case that non-white women are not less likely to be offered breast cancer directed surgery, rather that decision making regarding the perceived need for treatment or beliefs around treatment effectiveness varies by ethnicity. This may lead to non-white women declining breast cancer directed surgery, even when offered. Treatment beliefs have been highlighted as an ethnic barrier to accessing mental health services

e.g. (454). It is plausible that ethnic variations in cancer stigma and treatment perceptions are relevant too. Further work into the causes of ethnic variation in treatment utilisation outcomes are clearly still needed.

3.7.3 Strengths and Limitations

This study is the first to report English population-registry based study analysing utilisation of a high-cost targeted treatment (trastuzumab) in a publicly funded healthcare system. It is also one of the first studies to explore socio-economics inequalities in an emerging big data resource (SACT) as well as being the second largest population-based data analysis of trastuzumab utilisation after Du *et al.* (2011) (286).

There were limitations on account of the methods used. It is possible that the associations with treatment utilisation observed in this study reflect artefacts of the data collection process. However, the conclusion that the associations observed here occurred by chance seem unlikely given the variation in the magnitude of inequality observed across the range of HER2+ treatments (both conventional and novel) and the persistence of results even when the cohort was further restricted in sensitivity analyses.

The measure of SES is a limitation. IMD (as measured by the income domain) is an inference of SES and not necessarily the best way to determine this patient characteristic. As first detailed in Chapter 2, care is needed when interpreting the results to avoid the ecological fallacy (347). Whilst there are associations of reduced treatment utilisation at the group level (residence in an area of low deprivation), it is not possible to make the inference that individuals with a low SES experience reduced treatment utilisation. Additionally, as only one validated measure of SES (IMD's income domain) was explored in this analysis, it is not possible to conclude that utilisation associations would be the same with other SES measures e.g. level of education, occupation status. This will also mean that comparison of this analysis to other healthcare data using a different SES measure will be challenging. An in-depth review of SES measures is provided in Chapter 5, and this further discusses many of the points made here in greater detail across the breadth of the work in this thesis.

The need to use HER2 status to define the cohort of interest is a further potential limitation of the methods. This is because it may have resulted in some patients being excluded from this analysis who did in fact have a HER2+ breast cancer - but this was never identified because the tumour was never biomarker tested. This may have led to a smaller cohort which could be

biased, depending on the pattern of the missing HER2 status data. However, there is emerging evidence to suggest that uptake of HER2 testing in the UK is very high (455), so arguably the impact and potential for bias of refining a cohort based on a HER2 status variable is likely small.

There are also limitations of the SACT data source. First, the quality of early SACT data, prior to mandated trust submission post April 2014 (if not for longer given nationally discrepancies in electronic prescribing systems uptake), is known to be problematic (401) and the completeness of early SACT should therefore be cautioned (392,411). Inclusion of SACT data prior to April 2014 in these analyses may mean that the results do not fully reflect all trastuzumab prescribing at the time of data collection as trastuzumab utilisation in the early time frame of this analysis is likely underestimated. However, in sensitivity analyses exploring the role of using only more recent data (expected to be more complete) deprivation associations did not observe much deviation from the main model's findings. Second, there were many patients without SACT information recorded. Whilst it is possible to assume that a lack of a SACT reporting means that such patients did not utilise trastuzumab, this may not necessarily be the case. When considering the impact of this limitation on the findings of this study, it seems unlikely that missing SACT data was selectively biased by deprivation in so much that this would alter the direction of the effects observed. There was high comparability between patient demographics of those with and without SACT details (Appendix 3.6) and it is hard to envisage how mis-recording of utilisation by hospital administrative staff would be socio-economically biased. It is however possible that assumptions that missing SACT data reporting indicated trastuzumab non-utilisation is false. As these patients were included in analyses, this may again mean that overall utilisation percentages obtained here overall underreport true trastuzumab prescribing. Future analyses of NCRD and linked SACT data would benefit from assessment of impact of "missing" data with statistical methods such as multiple imputation.

Third, SACT data currently lacks details on drug indication. It is possible that there was some misclassification of trastuzumab utilisation for breast cancer, when in fact it was for gastric cancer. A restricted time frame of interest (trastuzumab prescribing 56 days prior to and up to 365 post diagnosis) along with selection of the cohort based on ICD code should minimise the likelihood that this was the case and there are probably not many women with HER2+ breast cancer in this time frame who also had a prior gastric cancer. Fourth, not all factors that may serve to act as a barrier to treatment utilisation are currently reported in SACT (or NCRD) e.g. frailty. Had such covariates have been available for inclusion in multivariable models, the

associations with deprivation observed here may have been attenuated. Another covariate of interest to this study - performance status - was requested as part of the initial ODR data application. However, upon receipt, it was evident that this variable was recorded in the SACT dataset and not the NCRD and referred the performance status at the start of the drug regimen (see Appendix 3.2 for the data dictionary at the time of data request). The main problem with this data field was the substantial missing data (n = 23,318; 57.79%) and that this was missing for those who did not have a systemic drug treatment. Additionally, the supplied variable's coding (0-9) did not match with either the standard Eastern Cooperative Oncology Group (ECOG) scale (0 [healthy] to 5 [dead]) (456) or Karnofsky Performance Status (KPS) scale (100 [no evidence of disease] to 0 [dead]) (457) which made classification somewhat challenging. This meant that it was not feasible to include this variable in multivariable models. This was an important limitation as performance status has a role in treatment decision making (e.g. precluding those in poor physical condition from accessing aggressive treatments) (458). Performance status is however (in theory) now available from the NCRD (ECOG scaling, although it should be noted that the completeness of this variable is unclear); hence future work exploring treatment utilisation should consider obtaining this variable to explore whether it may add to the analyses detailed here. Despite these limitations, the overall quality of SACT data has improved over time and the database is useful for providing SACT administration detail which was previously lacking with historic NCRD data (411).

Comorbidity data reporting is a limitation of the NCRD. This is because the CCM (437) is a crude measure of the number of comorbidities that a patient may have. The operationalisation of how comorbidities are recorded in cancer registries is an issue and as this data are derived from HES, comorbidities will reflect only those documented during an in-patient hospital admission. Hence under ascertainment is likely, especially for those patients with comorbidities diagnosed in primary care, those never admitted to hospital or those admitted to hospital who exhibited a comorbidity not listed (104,438). This may mean that the results are subject to residual confounding of comorbidities despite efforts being made to address this e.g. consideration of other factors which too may determine "fitness to treat" such as age. This limitation is discussed further in Chapter 5.

Generalisability of the study also needs consideration. This is because the study only reports a snapshot of trastuzumab prescribing for a set time in one country. Additionally, the study has only explored one novel anti-cancer therapy (and established treatment with clear predictive biomarker testing processes in one cancer sub-type - HER2+). Whilst there are other anti-

HER2+ novel anti-cancer therapies recorded in SACT (e.g. lapatinib, pertuzumab and trastuzumab emtansine), examination of the utilisation of such treatments was considered outside the remit of this work. In part this was due to time constraints but also the fact that overall utilisation numbers of these therapies in SACT were lower. This does mean though that the results obtained may not be generalisable to all other novel anti-cancer treatments in other countries where barriers (e.g. testing availability) and facilitators (e.g. clear clinical guidance) for treatment access may be vary along with the demographics of the cohort in question (greater ethnic diversity than reported here – denominator population was 88% white). Inequitable ethnic representation in genomic datasets is a well-known issue with sampling bias challenges – all of which can translate into clinical precision medicine care bias in diverse populations, where our understanding of gene-disease relationships may have been missed (459,460).

3.8 Conclusions

To conclude, this Chapter provides the first English analysis of trastuzumab utilisation in HER2+ breast cancer as reported in SACT data. Reduced treatment utilisation in patients residing in areas of greater deprivation was observed across all treatments examined though inequalities in surgery utilisation in contrast to trastuzumab utilisation were found to be more pronounced. Chemotherapy data was more complex. Overall, this Chapter found that despite advances in HER2+ breast cancer treatment, SES has a persistent role in determining treatment utilisation even in a publicly funded healthcare system. Further work now needs to explore whether inequalities with socio-economic status are also observed in other cancers, including those with less established novel anti-cancer treatments and predictive biomarker testing guidelines. An example of such a cancer is NSCLC - and analysis of treatment utilisation of a range of novel NSCLC anti-cancer therapies forms the basis of the next Chapter.

Chapter 4. Socio-economic Inequalities in NSCLC Cancer Treatment During the Era of Tumour Biomarker Guided Therapy: A Population-Based Study

4.1 Introduction

In Chapter 2 socio-economic inequalities in novel anti-cancer therapy were found to vary by cancer type and were larger in magnitude for lung than breast cancers. Chapter 3 advanced this understanding by exploring socio-economic inequalities in the utilisation of one novel anti-cancer therapy (trastuzumab) in a HER2+ breast cancer cohort in England. This population-based study found that a low SES was associated with reduced utilisation of trastuzumab and at a similar magnitude to that reported in Chapter 2. NSCLC provides many contrasts to HER2+ breast cancer patients, in terms of the number of novel anti-cancer therapies in use, the amount of time since their licensing, how well established these therapies and their associated predictive biomarker tests are within clinical guidance along with differences in the spread of incidence with SES quintiles. This final empirical Chapter therefore seeks to explore novel anti-cancer therapy utilisation in the context of a NSCLC population to establish whether in a publicly funded system, socio-economic inequalities in this cancer type are present and if so, what the magnitude of the effect is.

Chapter Aim: Determine the association of SES (measured in terms of the IMD income domain of area of residence at diagnosis) on novel anti-cancer therapy utilisation for a NSCLC population in England.

Objectives:

1) Conduct a retrospective observational cohort study using NCRD linked with the SACT dataset.
2) Undertake a primary analysis to determine the likelihood of utilisation of any novel anti-cancer therapy by SES for: (i) all patients; and (ii) stage IV patients.
3) Perform an exploratory analysis to determine the likelihood of utilisation of sub-groups of novel anti-cancer therapies (targeted therapy, EGFRi, (ALK inhibitors (ALKi), biologicals, and immunotherapy) by SES.

This Chapter begins by detailing the background to this study, before providing a summary of the steps taken to acquire and prepare the merged NCRD and SACT data for analysis. Results of the analysis then follow, along with a discussion of these within the context of the thesis research question. The Chapter ends with concluding comments.

4.2 Background

The move towards stratified care of selected populations with similar tumour biology is well exemplified in NSCLC where an evolving understanding of tumour heterogeneity, identification of novel biomarkers and the wealth of novel anti-cancer therapies (e.g. erlotinib, alectinib, atezolizumab) available has changed clinical practice in the last two decades (199). Many novel anti-cancer therapies now exist, including those indicated for stage IV disease, and these treatments have been shown to improve patient outcomes (72). Despite these medical advances, socio-economic gradients in NSCLC are historically well documented. A low SES is associated with: reduced conventional cancer treatment utilisation (221); increased mortality rate (103); and decreased survival rate (154). Chapter 2 provided evidence that globally, the uptake of NSCLC novel anti-cancer therapies is also subject to stark socio-economic patterning when compared to other cancers e.g. breast. Most of the published literature pertains to studies exploring healthcare systems such as the USA where there is a significant private component. Utilisation of these therapies at the English population level in a free at the point of delivery healthcare system, remains unknown. This retrospective cohort study therefore sought to determine whether there are socio-economic inequalities in the utilisation of NSCLC novel anti-cancer therapies in England.

4.3 Materials & Methods

4.3.1 Data Sources

This retrospective population-based cohort study used English NCRD linked with the SACT dataset. Details on both data sources have been previously discussed (see Chapter 3, Sections 3.3.1 and 3.3.2).

4.3.2 Ethics

Ethical approval for this study was obtained together with that for the HER2+ breast cancer cohort study (see Chapter 3, Section 3.3.4).

4.3.3 Office of Data Release Application

A joint application for release of breast and lung NCRD linked with SACT data was submitted to ODR and this has been previously described (see Chapter 3, Section 3.3.5).

4.3.4 Data Management: Lung File Preparation for Analysis

Data for all primary invasive lung and breast cancers diagnosed between 01/01/2012 - 31/12/2017 was provided in a single dataset. However, for the analysis in this Chapter, the cohort of interest were patients with NSCLC as this is the population for which there has been

great expansion in the range of precision medicines available. As the breast and lung data were supplied together, there was the potential that patients may have had multiple primary invasive lung tumours or tumours of both the breast and lung. To address this, initial data file management involved splitting the obtained data into two separate files for analysis by cancer type (one breast file and one lung). This process has already been discussed in detail (see Chapter 3, Section 3.3.6). To summarise, for this analysis, files were first reduced to provide a new file referring to only primary invasive tumours of the lung (n = 225,513; tumours/n = 228,115). Further cleaning (see Section 3.3.6, Step 3) then ensured that lung cancer data was reduced to one tumour per patient ID (see Section 3.3.6, Step 3). This was to deal with instance where registry data can capture multiple primary tumours for the same patient (e.g. more than one breast or lung tumour)[10]. After this reduction, the master lung data file (n = 225,513 tumours/n = 225,513 patients) resulted, and this was reduced to produce an analytical cohort pertaining to one tumour record per patient ID (n = 195,387 tumours/n = 195,387 patients). To achieve this, several inclusion and exclusion criteria were applied (Figure 4.1; page 151) and these are described below in more detail.

In the first instance, the datafile was reduced to retain only NSCLC cancers. This was achieved using the supplied codes for morphology (based primarily on ICD-O-02, though there are some instances where ICD-O-03 codes are referenced e.g. 8046). There were 89 morphology codes listed within the lung dataset which were categorised as follows as part of the data preparation:

- SCLC
- Other specified and non-specified
- NSCLC 8046 code (added in 2001 to ICD-O-03 to group cases that could not be classified beyond the exclusion of small-cell) (461)
- NSCLC Adenocarcinoma
- NSCLC Squamous
- NSCLC Large cell
- NSCLC 8000 code (malignant neoplasms not otherwise specified (NOS))
- NSCLC 8010 code (carcinoma in situ, NOS)

This categorisation was chosen as based clinical input from a supervisor (AG)[11], along with a range of sources describing lung cancer sub-types differentiation based on morphology codes (461–466).

[10]Individual patients could be in both breast and lung ODR files if they had both a breast and lung cancer diagnosed in the study period.
[11]Refers to Dr. Alastair Greystoke.

Morphology codes which corresponded to "SCLC" (n = 23,180 tumours) and "other specified or unspecified" tumours (n = 6,944) were excluded (Figure 4.1). This left several groups of codes which fell under the umbrella term "NSCLC": adenocarcinomas (n = 69,110); squamous cell tumours (n = 41,534); large cell tumours (n = 1,791); and the 8000 (n = 38,412), 8010 (n = 28,892) and 8046 codes (n = 15,650). The codes 8000, 8010 and 8046 represent non-specific NSCLC classifications. These were grouped together as "Not otherwise specified (NOS) NSCLC" and retained in the analysis dataset despite their non-specificity as, when combined, they captured large patient numbers. A further histological division of "non-squamous NSCLC" was also defined for use in later sensitivity analyses. This grouping combined tumours only with the following morphology: an adenocarcinoma; NSCLC NOS; or large cell. This action was to taken to account for the fact that the majority of NSCLC with a mutated oncogenic driver occur within the adenocarcinoma histology, hence clinical guidelines recommend predictive biomarker testing in the non-squamous population for this reason (467,468).

Of the 195,389 patients with a first primary invasive NSCLC, one patient was excluded as they had no IMD status recorded so no associations with deprivation could be determined. A further patient was excluded as their tumour was classified as stage 0 and it was deemed unlikely that novel anti-cancer therapies would be utilised. The decision was taken not exclude early stage NSCLC, despite the fact that during the time frame of this analysis, novel anti-cancer therapies were not usually indicated in such patient groups. The rationale for this decision was because the registry records the stage of the disease at diagnosis (rather than stage of disease at treatment) so it was possible that the patients' disease progressed between time of diagnosis and treatment. Also, as overall utilisation was low, efforts were made to include all incidences of novel anti-cancer therapy use (Appendix 4.1) within this analysis where possible. At the end of this process, a denominator analytical population of 195,387 patients resulted (Figure 4.1).

Of the denominator population, 145,343 patients had no linked SACT record. For those patients who did have an associated SACT record (n = 50,044), further data cleaning was required as there can be multiple drug entries in SACT per patient ID depending on how many and how often patients receive each SACT drug treatment. This analysis was only concerned with whether a patient utilised a novel anti-cancer therapy (including utilisation of each novel anti-cancer therapy sub-group e.g. targeted therapy, immunotherapy and biologicals) and not how many times, they utilised treatment(s). To assist with data file cleaning to achieve this and to increase the likelihood that these SACT treatments were used for the primary invasive NSCLC cancer of interest, a time frame restriction was applied. Consideration was restricted to SACT

treatments administered within 56 days prior to diagnosis and up to 2 years after the diagnosis of the incident NSCLC cancer. This period was wider for that used in the breast analysis to account for prescribing differences between the two cancer types. In NSCLC, it is likely that a novel anti-cancer therapy could be started sometime after diagnosis, including as a second or third-line treatment. Based on clinical advice and considering the rapid prognosis of lung cancer, a two-year time frame post-diagnosis was considered reasonable. As per the breast analysis, a limit of 56 days prior to diagnosis was also chosen to allow for differences in how "data of incidence" is recorded on NCRD (which follows a standard algorithm) and how the point of diagnosis might be defined in clinical practice. At the end of this part of the data cleaning process, there were 41,998 NSCLC patients with a SACT drug entry within the desired time frame out of a total analytical cohort of 195,387 patients.

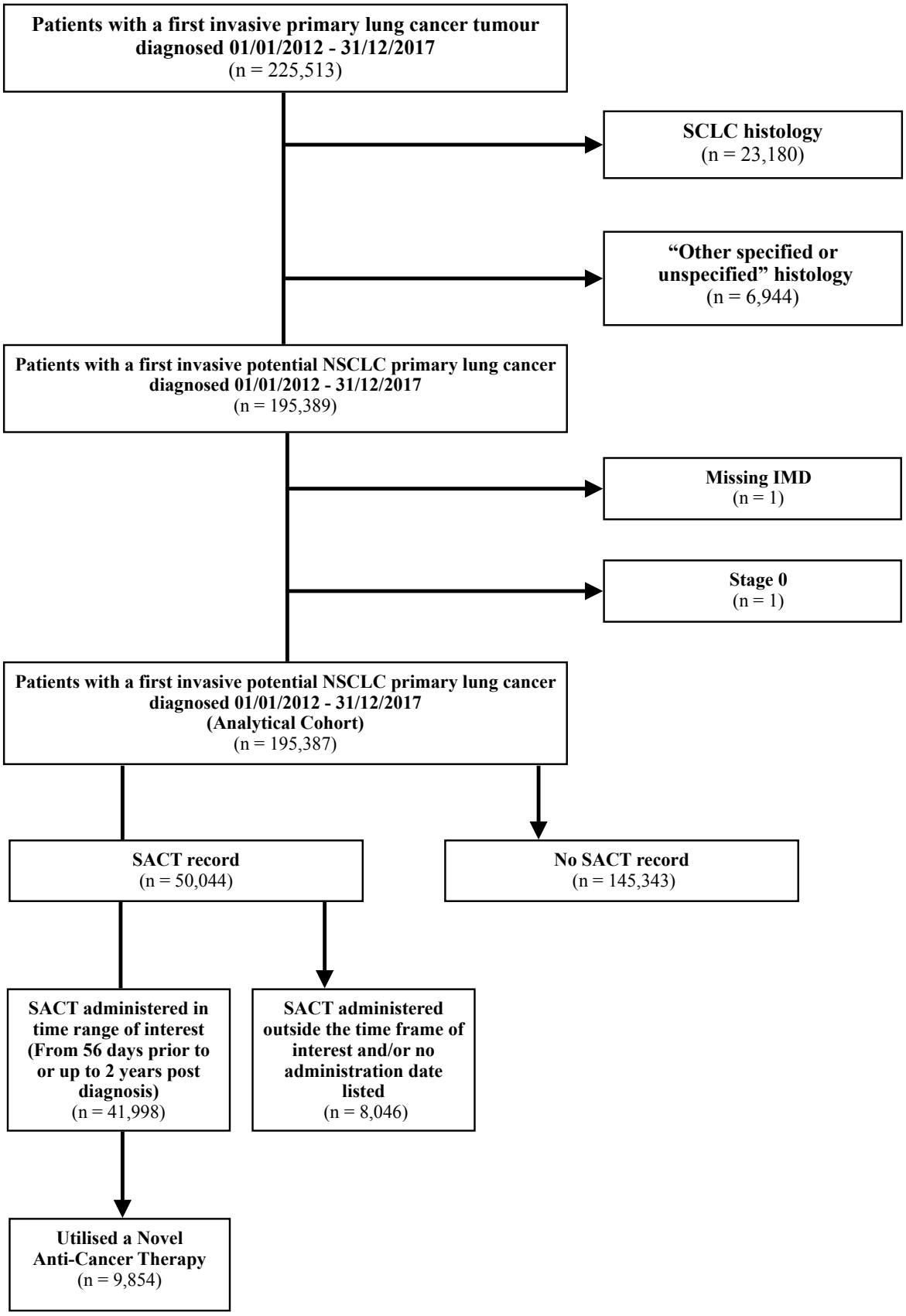


Figure 4.1 Flow diagram depicting the analytical cohort.

4.3.5 Missing and Unknown Data

As per Chapter 3, missing data in the cancer registry was also a feature in this analysis. Methods and rationales for dealing with missing and unknown data have been previously described (see Section 3.3.8). Table 4.1 outlines the missing data in this analysis, along with strategies used to account for this.

Table 4.1 Missing, unknown, and other data in the NSCLC dataset. Abbreviations: IMD: Index of multiple deprivation; SACT: Systemic anti-cancer therapy; SES: socio-economic status.

Variable	Missing/ Unknown Data	Data Handling Strategy
No SACT Record	145,343 (74.39)	Included in the analytical cohort as SACT record only generated if a patient has systemic anti-cancer drug treatment. Characteristic compared to those of the group with a SACT record. Highlighted in limitations as a high proportion of unknown data.
Ethnicity	8,690 (5.45)	Labelled as "missing/unknown". Included in multivariable analyses.
Stage	320,031 (10.25)	Labelled as "unknown". Included in multivariable analyses or analyses split by staging (stage IV only). Unlikely missing as random – likely mainly comprises older patients with more advanced disease (and with other comorbidities) who were not sent for staging.
IMD	1 (0.00)	Excluded from cohort as small number and no association of novel anti-cancer therapy utilisation by SES could be determined.
Performance Status at Start of Drug Regimen	161,305 (82.56)	Unable to use in this analysis as substantial missing data (over 80% of patients) and only available for those with a SACT record.

4.4 Outcome Measures

Primary Analysis - Any Novel Anti-Cancer Therapy: The primary focus of this Chapter was to discern whether each patient received a novel anti-cancer therapy - of any type (targeted treatment, biological or immunotherapy). The rationale for combining treatments in this way was that the relatively short time window of the analysis, combined with the rarity of some tumour mutations (98), meant that it was necessary to combine therapies.

As SACT currently does not include information on drug indication, the term "novel anti-cancer therapy" was first defined.

Utilisation of any novel anti-cancer therapy was identified as follows:

- A targeted treatment (EGFRi, ALKi, or other sub-division), biological, immunotherapy drug(s), and/or other (novel anti-cancer therapy which did not fall under one of the other classes mentioned e.g. drugs listed only by a preclinical drug number and/or first in class drugs).
- A record of a novel anti-cancer therapy in any of the drug, benchmark, or analysis group data fields.
- A novel anti-cancer therapy record, either alone or in combination with other treatments (e.g. bevacizumab, carboplatin, and paclitaxel).
- A clinical trial where it was clear that all patients would have received the novel anti-cancer drug, regardless of the trial arm. As SACT only records clinical trials by name (rather than listing the drugs involved), the clinical information to aid this decision making was derived from trial design information such as trial databases such as CRUK clinical trials finder (424) or clinical trials.gov (425) if the trial was still active, from trial publications or through guidance provided from a lung oncologist (AG)[12].

This definition was compiled using the therapies listed in the denominator population records. Several steps were taken to minimise misclassification of drugs. First, NICE guidelines (194), the BNF (246), EMC (427), EMA (428), Northern Cancer Alliance protocols (429), ASCO (203), and ESMO guidelines (430) were used to identify current and old licensing of the therapies of interest during 2012 - 2017. Second, clinical advice was sought (AG)[9], especially regarding pre-licensed drugs. Finally, decisions erred on the side of caution and therapies were included where it was likely, given ICD codes and time frame restrictions on the dataset, that drug prescriptions would be for a NSCLC indication. Novel anti-cancer therapies with no NSCLC indication were excluded.

Of the 344 analysis groups, 327 benchmark groups and 186 drug groups supplied as part of the ODR data request, only 33 analysis groups, 33 benchmark groups and 41 drug groups were included as meeting the definition “utilisation of novel NSCLC anti-cancer therapy” (Appendix 4.2). Next, for each drug/SACT entry per patient ID, a binary variable was created (presence or absence of a SACT therapy listed in Appendix 4.2). This was because this analysis was only interested in utilisation of any novel anti-cancer therapy and not the number of novel anti-cancer therapies prescribed. Duplicate references to novel anti-cancer entries were dropped by patient ID (the first instance of each record per patient ID was kept). The left only one drug record per patient ID (n = 195,387).

[12]Refers to Dr. Alastair Greystoke.

The primary analysis was conducted for all NSCLC stages to explore drug utilisation across the whole population. This was considered important to help maximise overall novel anti-cancer therapy numbers given the low anticipated utilisation rate for such treatments in rarer biomarker indications. However, as most NSCLC is diagnosed at a late stage and anti-cancer therapy has primarily benefited the stage IV cohort, a repeated analysis, restricting the population was stage IV only was also performed.

Exploratory Analysis - Targeted Therapy, Biologicals & Immunotherapy: For the exploratory analysis, which considered utilisation of different groups of novel anti-cancer therapies by SES (as measured by deprivation), the therapies were categorised further into the following six groups:

(i) Any Targeted Therapy: This refers to any novel anti-cancer therapy defined as a targeted treatment and includes EGFRi, ALKi and other kinase inhibitor drugs.

(ii) EGFRi: This refers to a specific type of targeted drug that inhibits EGFR receptor activity (e.g. gefitinib).

(iii) EGFRi not including erlotinib: This refers to any group (ii) drug but excludes erlotinib. This analysis was conducted as up to 2015, erlotinib had NICE approval in EGFR wild-type NSCLC and this may have resulted in some unselected population application. It was only after this date that a positive EGFR result became a prerequisite for access.
(IV) ALKi: This refers to drugs acting on variations in ALK such as EML4-ALK translocation (e.g. crizotinib).

(V) Biologicals: This refers to drugs, such anti-angiogenics, which ordinarily would support tumour growth (e.g. bevacizumab).

(VI) Immunotherapies: This refers to drugs which activate the immune system to elicit an immune response (e.g pembrolizumab).

The SACT drug, benchmark, and analysis codes for each of these six groupings are outlined in Appendix 4.3. As per the primary analysis outcome, these classifications were allocated based on information pooled from NICE guidelines (194), the BNF (246), EMC (427), EMA (428),

Northern Cancer Alliance protocols (429), ASCO (203), and ESMO (202) guidelines as well as with discussion with AG[13]. Analyses were run for all stages of disease combined.

4.4.1 Main Explanatory Measure: Deprivation

As per the breast analysis, IMD of the patient's area of residence at the time of cancer diagnosis was used as a relative measure of SES. This measures deprivation using quintile subdivisions derived from the income domain of IMD. Chapter 3 (see Section 3.4.1) provides further information on the NCRD IMD variable. As the dataset covered the time period 2012 to 2017, NCRAS provided two IMD measures (referring to 2010 and 2015). Thus, for cancers diagnosed in 2012, the IMD 2010 variable was used to define deprivation, while for cancers diagnosed in 2013, 2014, 2015, 2016 and 2017, the IMD 2015 measure was used. This ensured that the deprivation status used in the analysis represented that to the closest IMD measure at the time of cancer diagnosis.

4.4.2 Other Explanatory Measures[14]

Sex: Sex was coded as male and female.

Age: Diagnosis age for lung cancer patients was organised into six groups (<50, 50-59, 60-69, 70-79, 80-89 and 90+).

Year of Diagnosis: Period of diagnosis was determined by diagnosis year. As per the breast analysis, time was also an important factor for consideration in this Chapter's sensitivity analyses. This is because changes in SACT recording following mandatory SACT submission post April 2014 may have impacted the quantity of SACT lung cancer drug capture prior to this date so early SACT data may not accurately reflect the full population utilising treatment at that time.

Ethnicity: Ethnicity was grouped as: white; other ethnic group (Asian/British, Asian, Black/African/Caribbean/Black British, mixed/multiple ethnic groups and other ethnic groups); and missing/unknown (refers to missing and unknown ethnicity classifications). As per the breast analysis, lung cancer patients' ethnicity was primarily white (n = 179,690; 91.97%) and other individual ethnic groups were rare, hence why these classifications were chosen.

[13]Refers to Dr. Alastair Greystoke.
[14]Not all explanatory measures of interest were available at the time of data request. This limitation is discussed further in the discussion (see Section 4.7.3).

Rural/Urban Indicator: Rural/urban indicator was categorised as: rural village (including in a sparse setting); hamlet and isolated dwellings (including in a sparse setting); rural town and fringe (including in a sparse setting); urban city and town (including in a sparse setting); urban conurbation (minor and major). A more detailed description of how this variable is determined is provided in Chapter 3 (see Section 3.4.2).

Stage: Stage was grouped as I, II, III, IV and unknown (stage unknown or unstageable tumours). As lung cancer is often diagnosed at a late stage, a restricted stage IV cohort analysis was also undertaken as this population represents a significant proportion of the dataset (n = 90,785; 46.5%) which may be important in determining treatment utilisation.

Multiple Tumours: This variable, called "big tumour count' in the NCRD was described in detail in Chapter 3 (see Section 3.4.2) and refers to tumours other than the index lung cancer. For this analysis, classification was simplified into two groups: one primary tumour and more than one primary tumour.

Comorbidities: This variable was grouped as 0, 1-2, or 3+ comorbidities diagnosed between 78 to 6 months prior to diagnosis. Further details on the specifics of the algorithm used by NCRAS to generate this variable is described in Chapter 3 (see Section 3.4.2).

4.5 Statistical Analysis

The purpose of the analysis was two-fold. A primary analysis aimed to determine whether there are socio-economic associations in the utilisation of any novel anti-cancer therapy for: (i) all patients (of any and unknown stage); and (ii) in secondary analysis, all stage IV patients. A subsequent exploratory analysis then explored socio-economic associations in novel anti-cancer therapy utilisation and whether these varied by therapy sub-group (any targeted therapy, EGFRi, EGFRi but not including erlotinib, ALKi, biologicals, and immunotherapies).

Baseline cohort demographic and clinical characteristics (number and percentage) were reported for the whole cohort, stage IV patients, and those without a SACT record. For any novel anti-cancer therapy, utilisation descriptive statistics (number and percentage) were provided by all independent variables of interest. Chi-square tests were used to determine whether there was an association between the demographic/clinical characteristics and treatment utilisation. Additional chi-square tests were then used to test for associations between deprivation and all other explanatory variables.

Logistic regression models were then developed to determine associations between SES and: (i) any novel anti-cancer therapy; and (ii) any novel anti-cancer therapy in stage IV patients. Explanatory clinical and demographic variables were considered for inclusion in the multivariable models if they were significant (LRT $p \leq 0.05$) in univariable analyses. Deprivation as the primary variable of interest, was forced into all multivariable models. Models were reduced so that they contained only variables which remained statistically significant (LRT $p \leq 0.05$) in the presence of other variables. BIC and AIC were used to inform choices between alternative models. Model goodness of fit was determined using the Hosmer and Lemeshow χ^2 tests. Additional checks for collinearity included ensuring that mean variance inflation factors were lower than 10 in the final models.

The primary analysis model (all stages) included deprivation and was adjusted for the following covariates: age; diagnosis year; ethnicity; rural/urban indicator; stage; multiple tumours; comorbidities and histology. The stage IV model included SES and was adjusted for age, diagnosis year, ethnicity, rural/urban indicator, multiple tumours, and comorbidities. Unadjusted and adjusted ORs are reported with 95% CI and p values for all models. In these models, deprivation/SES was fitted as a categorical variable. The models were re-run fitting it as a continuous variable to test for evidence of a linear trend across categories.

Utilisation of novel anti-cancer therapies in combination with other conventional treatments by IMD was explored but as treatment ordering is not considered important clinically in NSCLC and as the data showed no different patterns of association with deprivation in the primary analysis, this preliminary analysis was not pursued further.

Three sensitivity analyses were performed by limiting the full patient cohort (any stage) to: (i) adenocarcinoma histology (sensitivity analysis 1); (ii) non-squamous histology (sensitivity analysis 2); and (iii) date of incidence from April 2014 onwards (sensitivity analysis 3). Analysis of adenocarcinomas was undertaken because this histological group would, clinically, be the most likely to receive these treatments. A non-squamous analysis was undertaken given targeted therapy utilisation for these patients was anticipated to be higher (likely to have molecular testing). Finally, the restriction on the dataset from April 2014 onwards was to allow for the fact that this was the date when SACT recording became mandatory for all Trusts.

When the different novel anti-cancer therapy sub-groups were considered, an exploratory, minimally adjusted analysis (rather than a fully adjusted multivariable analysis) was

undertaken. This was because as the numbers utilising novel anti-cancer therapies were expected to be small (i.e. utilisation was an uncommon outcome). For each drug grouping, a model was fitted and adjusted for sex, age, and ethnicity. These factors were deemed most relevant given their links to tumour biology; they provide a rationale for why a patient is more likely to have a targetable drug mutation. As per the primary analysis, patients of all disease stages were included and (minimally) adjusted ORs are reported along with 95% CI and p values. In these models, deprivation/SES was fitted as a categorical variable. The models were re-run fitting IMD as a continuous variable to test for evidence of a linear trend across categories.

All statistical analyses were conducted using STATA version 16.1 (StataCorp, College Station, TX).

4.6 Results

4.6.1 Cohort Characteristics

All Patients: There were 195,387 patients diagnosed with a first, invasive primary NSCLC between 01/01/2012 - 31/2/2017. Most of these patients were: aged between 60-85 years old (84.7%); of white ethnicity (92.0%); and resident in urban areas at the time of diagnosis (82.5%). Males made up over slightly half the cohort (54.1%) and almost half of patients had stage IV disease (46.5%). The largest proportions of patients presented with NSCLC NOS (42.5%) or NSCLC adenocarcinomas histology (35.4%). 54.7% of patient had no comorbidities and 45.3% had or one or more. The percentage of patients' resident in each deprivation quintile was as follows: 14.1% for IMD 1 (least deprived);18.0% for IMD 2; 20.0% for IMD 3; 22.4% for IMD 4; and 25.5% IMD 5 (most deprived). The number of primary invasive lung NSCLC cases diagnosed each year remained consistent over the period of the data covered (16.4% of total NSCLC were diagnosed in 2012 and 16.7% in 2017) (Table 4.2).

Table 4.2 Demographic and clinical characteristics of the cohort with a first invasive primary NSCLC diagnosed between 01/01/2012 – 31/12/2017 (n = 195,387)

	Number (%)
IMD[1]	
1 (Least Deprived)	27,634 (14.14)
2	35,198 (18.01)
3	39,007 (19.96)
4	43,660 (22.35)
5 (Most Deprived)	49,888 (25.53)
Sex	
Male	105,717 (54.11)
Female	89,670 (45.89)
Age (Years)	
<50	4,206 (2.15)
50 – 59	16,609 (8.50)
60 – 69	48,561 (24.85)
70 – 79	68,294 (34.95)
80 – 89	48,557 (24.85)
90+	9,160 (4.69)
Diagnosis Year	
2012	32,130 (16.44)
2013	32,428 (16.60)
2014	32,561 (16.66)
2015	32,669 (16.72)
2016	32,986 (16.88)
2017	32,613 (16.69)
Ethnicity	
White	179,690 (91.97)
Other Ethnic Group[2]	7,007 (3.59)
Missing/unknown[3]	8,690 (4.45)
Rural/Urban Indicator	
Rural Village, Hamlet & Isolated Dwellings	15,384 (7.87)
Rural Town & Fringe	18,798 (9.62)
Urban City & Town	86,505 (44.27)
Urban Conurbation	74,700 (38.23)
Stage	
I	32,408 (16.59)
II	15,321 (7.84)
III	36,842 (18.86)
IV	90,785 (46.46)
Unknown[4]	20,031 (10.25)

[1]For diagnosis year 2012, IMD_2010 was used; for diagnosis years 2013 -2017, IMD_2015 was used.
[2]Other ethnic group refers to Asian/British Asian, Black/African/Caribbean/Black British, mixed/multiple ethnic groups and other ethnic groups.
[3]Missing/unknown refers to unknown and missing ethnicity classifications.
[4]Unknown staging refers to missing and unstageable tumours.

Table 4.2 Continued

	Number (%)
Histology	
NSCLC NOS	82,953 (42.46)
NSCLC Adenocarcinoma	69,109 (35.37)
NSCLC Squamous	41,534 (21.26)
NSCLC Large Cell	1,791 (0.92)
Multiple Tumours	
1	159,130 (81.44)
>1	36,257 (18.56)
Number of Comorbidities (Between 78 to 6 Months Prior to Diagnosis)[5]	
0	106,870 (54.70)
1 – 2	61,024 (31.23)
3+	27,493 (14.07)
SACT Record	
Yes (In Time Range)[6]	41,998 (21.49)
Yes (Not in Time Range)[6]	8,046 (4.12)
No	145,343 (74.39)
Treatment	
Utilised Chemotherapy	49,273 (25.22)
Utilised Surgery	29,839 (15.27)
Utilised Radiotherapy	53,835 (27.55)
Utilised Any Novel Anti-Cancer Therapy	9,854 (5.04)
Utilised a Targeted Therapy	4,783 (2.45)
Utilised a Biological	1,039 (0.53)
Utilised an Immunotherapy	4,398 (2.25)

[5]Measured using the Charlson Comorbidity Index
[6]Refers to between 56 days prior to or up to 2 years days post diagnosis.
Abbreviations: IMD: Index of Multiple Deprivation; NSCLC: non-small cell lung cancer; NOS: not otherwise specified; SACT: Systemic Anti-Cancer Therapy Database.

Stage IV: When restricting the dataset to stage IV NSCLC cases only (n = 90,785; 46.5% of full cohort), clinical and demographics were comparable to those of the main cohort. Most patients were: aged between 60–85 (83.3%); of white ethnicity (91.9%); and resided in urban areas (82.1%), while males made up just over half of the cohort (55.1%). Many patients had no comorbidities (60.5%). There were slightly more patients diagnosed with a NSCLC NOS morphology (43.3%) than adenocarcinoma (40.3%) when compared to the full cohort. Diagnoses per year remained consistent (16.5% of the total diagnoses in 2012 compared to 16.5% in 2017). The percentage of patients in each deprivation quintile was as follows: 14.5% for IMD 1 (least deprived); 18.3% for IMD 2; 20.3% for IMD 3; 22.2% for IMD 4; and 24.7% for IMD 5 (most deprived) (Appendix 4.4).

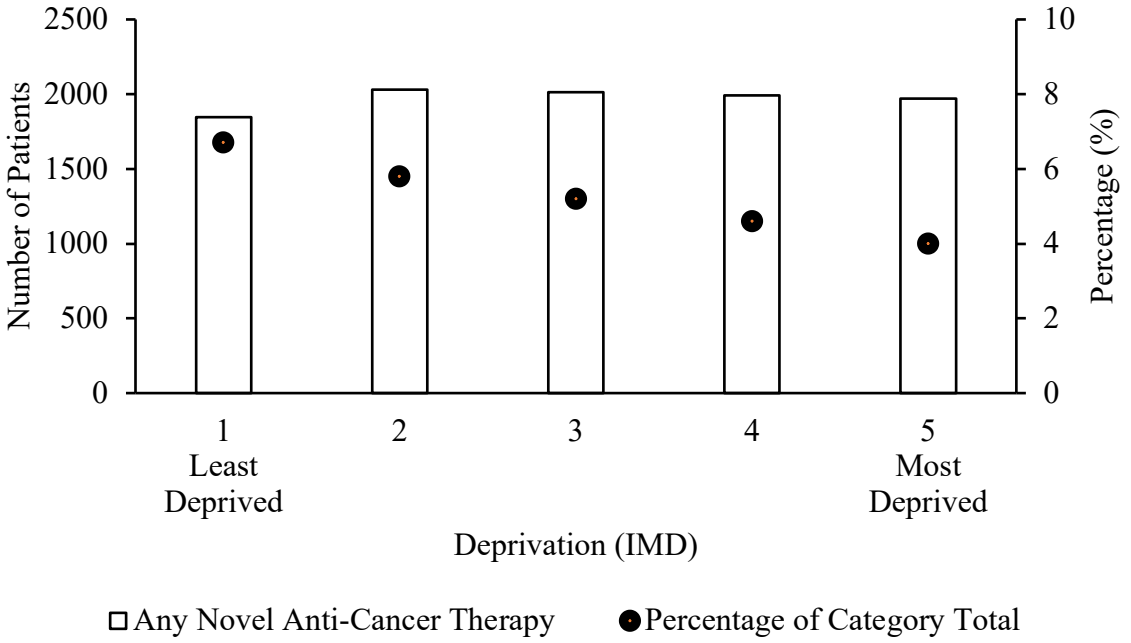
SACT Record in Time Range: A significant proportion of patients in the full cohort had no SACT record (74.4%). For patients who did have a SACT record and within the time frame of interest (n = 41,998), similar trends in demographics and clinical characteristics when comparing to patients without a SACT record (n = 145,345) were observed (Appendix 4.5). The number of patients with a linked SACT record did increase over the frame of the analysis (in 2012, 21.1% of total records had a SACT record and 27.5% by 2017).

Demographic and Clinical Characteristics by Deprivation Category: Data for the full dataset are shown in Appendix 4.6. Patients' resident in the more affluent areas were less often of non-white ethnicity than those resident in the most deprived areas. The least deprived areas had a lower fraction of non-white patients. They were more likely to reside in a rural setting than those of higher deprivation and were more likely to have an NSCLC adenocarcinoma histology.

4.6.2 Primary Analysis: Any Novel Anti-Cancer Therapy (All Patients)

Descriptive Statistics: 5% (n = 9,854) NSCLC patients of any stage utilised a novel anti-cancer therapy. Utilisation was higher in patients when resident in the least deprived areas (6.7%) compared to those with resident in the most deprived areas (4.0%) (Figure 4.2a). Utilisation of novel anti-cancer therapies increased over time; from 2.7% among patients diagnosed in 2012 to 10.2% of those diagnosed in 2017 (Figure 4.2b). Stage III and IV NSCLC were more likely to receive novel anti-cancer therapies (combined n = 8,852; 4.5% of all NSCLC patients) (Figure 4.2c). Utilisation decreased by age (16.0% in the <50 age group compared with group 0.32% in the 90+ age group) (Figure 4.2d). Utilisation was slightly higher in females (5.6%) compared to males (4.6%; Figure 4.2e). The proportion of non-white patients utilising novel anti-cancer therapies was more than double that of white patients (11.7% vs 4.9%) (Figure 4.2f). Utilisation declined as the number of comorbidities increased. Most patients utilising novel anti-cancer therapies had no comorbidities (Figure 4.2g).

a)



b)

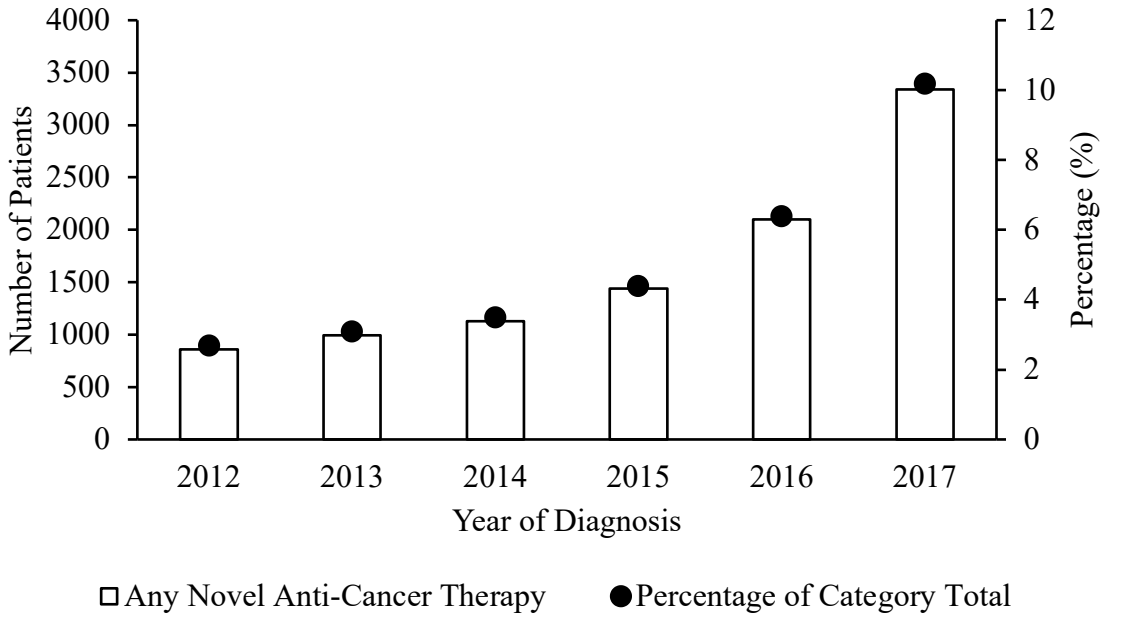
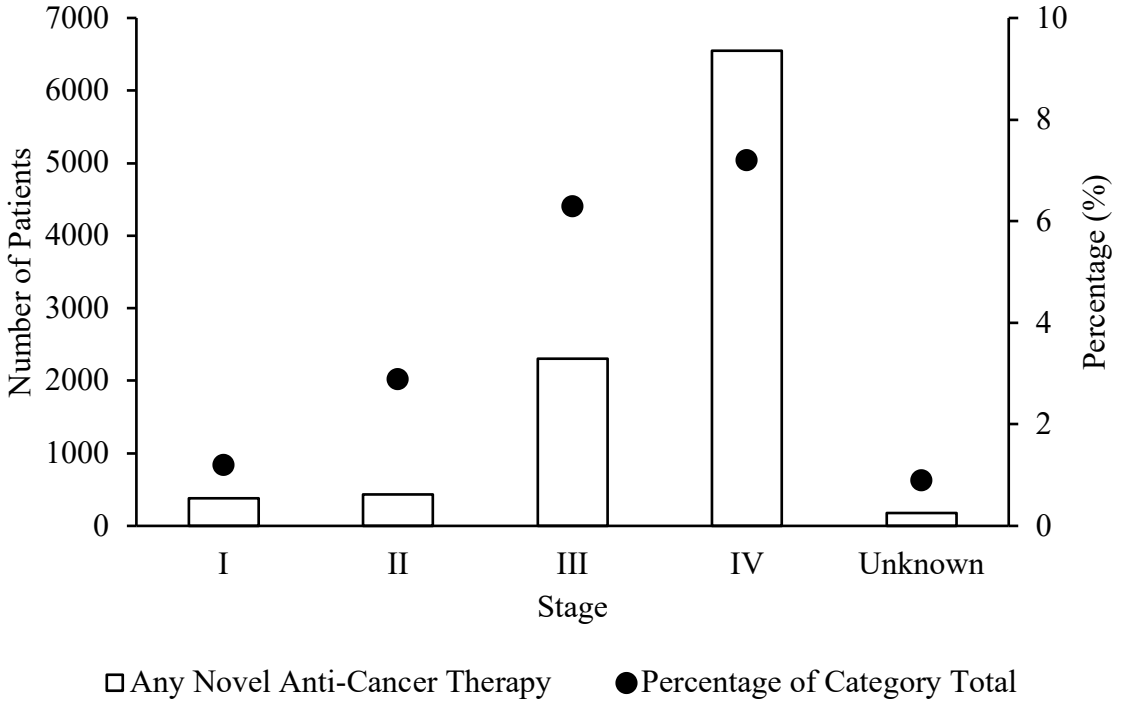


Figure 4.2 Any novel anti-cancer therapy utilisation by: **a)** each quintile of deprivation by IMD (IMD 1 = least deprived); and **b)** diagnosis year. IMD: Index of multiple deprivation.

c)



d)

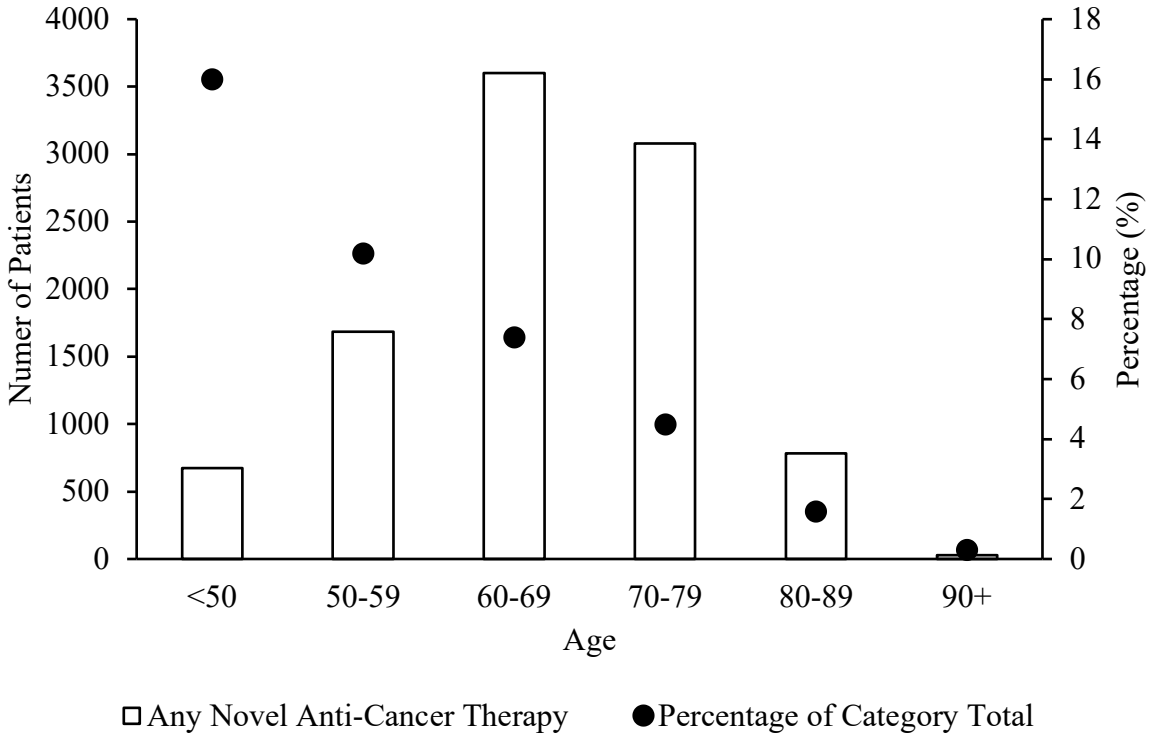
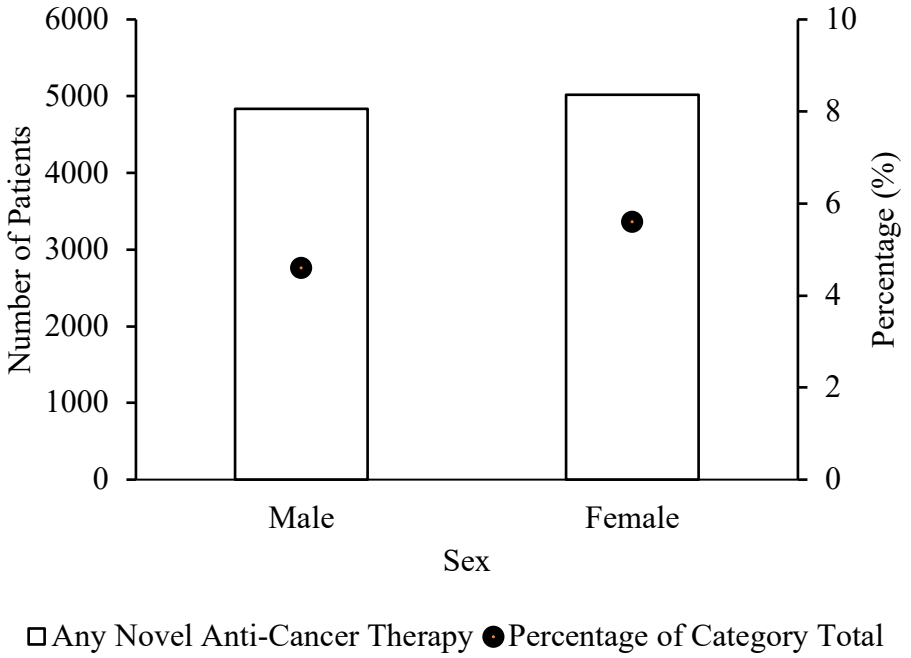


Figure 4.2 Any novel anti-cancer therapy utilisation by: **c)** stage; and **d)** age.

e)



f)

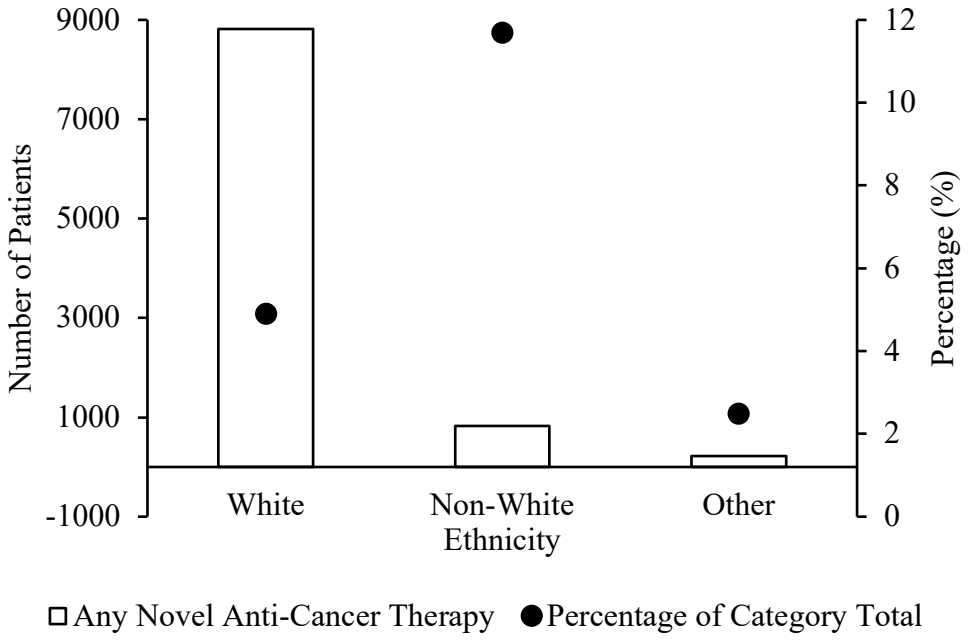


Figure 4.2 Any novel anti-cancer therapy utilisation by: **e)** sex; **f)** ethnicity.

g)

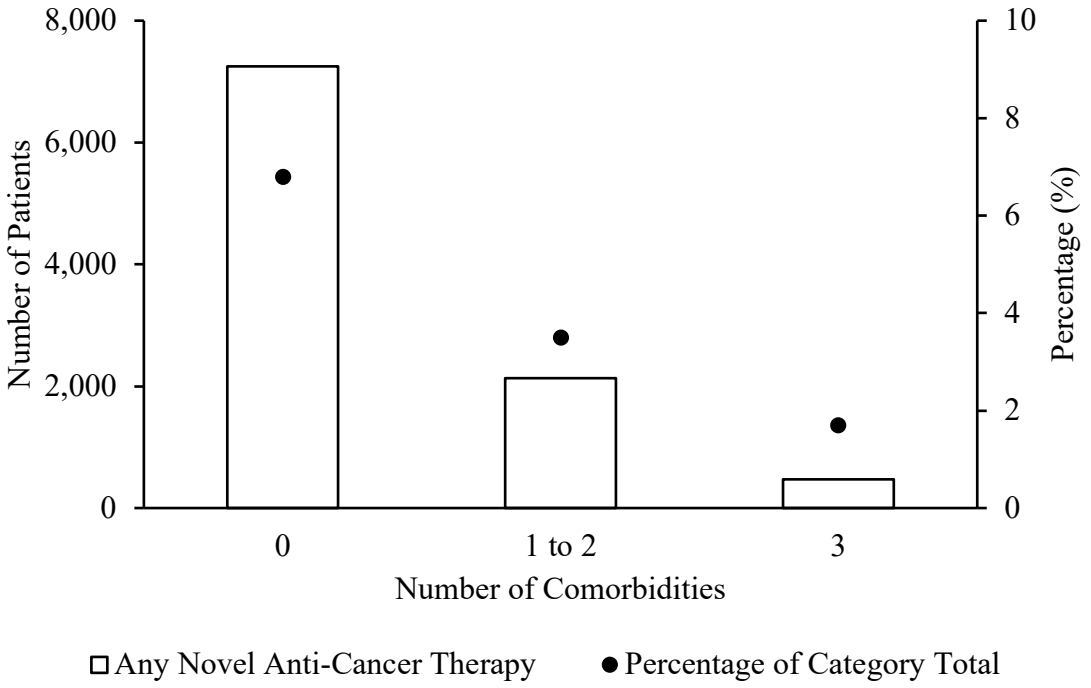


Figure 4.2 Any novel anti-cancer therapy utilisation by: **g)** number of comorbidities.

Multivariable Model Deprivation Associations: Utilisation of novel anti-cancer therapies reduced as the level of deprivation increased: 6.7% (IMD 1 least deprived); 5.8% (IMD 2); 5.2% (IMD 3); 4.6% (IMD 4); and 4.0% (IMD 5 most deprived). This trend for decreasing utilisation by deprivation persisted after adjustment for confounders in the multivariable model. The final multivariable model showed a statistically significant association between any novel anti-cancer therapy receipt and deprivation (LRT $p \leq 0.05$). There was a pattern of decreasing odds of utilisation with increasing deprivation. Patients residing in the most deprived areas were 46% less likely to utilise novel anti-cancer therapies compared to those residing in the least deprived areas (IMD 5 vs IMD 1; mvOR 0.54, 95% CI 0.50, 0.58) (Table 4.3). The test for linear trend was significant ($p \leq 0.05$) (Figure 4.3).

Sensitivity Analyses: When restricting analyses to adenocarcinomas (n = 69,109; sensitivity analysis 1), 10.2% (n = 7,012 patients) were found to have utilised a novel anti-cancer therapy. A significant association of novel anti-cancer therapy utilisation with deprivation was observed (LRT $p \leq 0.05$) along with a trend of decreasing utilisation with increasing deprivation across each IMD quintile. A similar magnitude of reduced novel anti-cancer therapy utilisation was observed by deprivation category (IMD 5 vs IMD 1; mvOR 0.5l, 95% CI 0.51, 0.60) (Appendix

4.7). When restricting analyses to non-squamous histology (n = 153,853; sensitivity analysis 2) 5.3% (8,123 patients) utilised a novel anti-cancer therapy. A significant association of deprivation with novel anti-cancer therapy utilisation was observed (LRT $p \leq 0.05$). There was also a trend across IMD quintiles of decreasing utilisation with increasing deprivation. A stronger reduction in utilisation by deprivation was also observed across all deprivation categories as well as when comparing the least deprived and most deprived residents (IMD 5 vs IMD 1; mvOR 0.46, 95% CI 0.43, 0.50) (Appendix 4.8). Finally, following restriction of the cohort to cancers diagnosed after mandatory SACT submission (April 2014) (n = 122,708; sensitivity analysis 3), 6.3% (n = 7,717 patients) utilised a novel anti-cancer therapy. Overall, the association of deprivation with novel anti-cancer therapy utilisation was significant (LRT $p \leq 0.05$) and again a trend was observed across the IMD quintiles (decreasing novel anti-cancer therapy utilisation with increasing deprivation). In this analysis similar patterns to the primary analysis were seen (IMD 5 vs IMD 1; mvOR 0.53, 95% CI 0.49, 0.58) (Appendix 4.9).

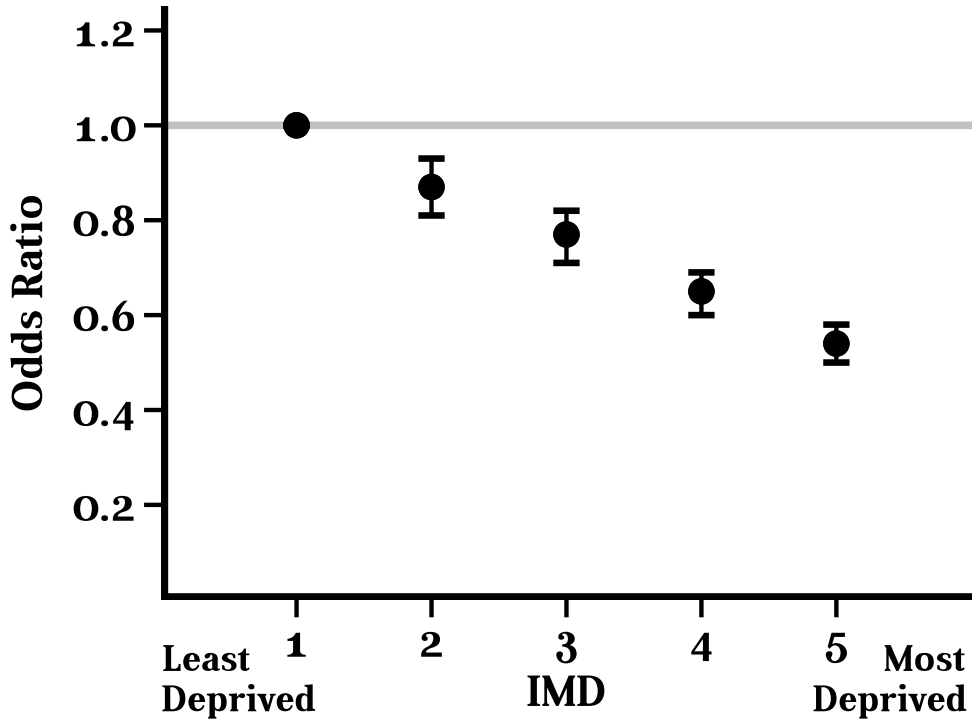


Figure 4.3 Any novel anti-cancer therapy utilisation by deprivation (IMD 1 least deprived; IMD 5 most deprived). Multivariate model adjusted for: sex, age, diagnosis year, ethnicity, rural/urban indicator, stage, comorbidities, multiple tumours, and histology. IMD: Index of multiple deprivation.

Multivariable Model Other Variables & their Associations with Utilisation: Several variables in the multivariable model, other than deprivation, also had statistically significant associations with reduced any novel anti-cancer therapy utilisation. These were: male sex (male vs female; mvOR 0.81, 95% CI 0.78, 0.85); older age (80-89 vs 70-79; 0.47, 95% CI 0.43, 0.51); white ethnicity (other ethnic group vs white mvOR; 1.97, 95% CI 1.81, 2.15); a lower staged tumour (stage I vs stage IV; mvOR; 0.13, 95% CI 0.12, 0.15); and multiple comorbidities (3+ vs 0 comorbidities mvOR; 0.45, 95% CI 0.41, 0.50).

Table 4.3 Likelihood (OR and 95% CI and p values from logistic regression) of utilising any novel anti-cancer therapy (targeted therapy, immunotherapy, or biologic) by deprivation and adjusted for: age, diagnosis year, ethnicity, rural/urban indicator, stage, multiple tumours, comorbidities, and histology for patients with a NSCLC diagnosed between 01/01/2012 - 31/12/2017 (n = 195,387)

				Unadjusted			**Adjusted**		
	Number (%) Utilising a Novel Therapy n = 9,854 (5.04)	Number (%) Not Utilising a Novel Therapy n = 185,533 (94.96)	P Value[1]	OR	95% CI	P Value[2]	OR	95% CI	P Value[2]
IMD[3]			<0.001			**<0.001**			**<0.001**
1 (Least Deprived)	1,847 (6.68)	25,787 (93.32)		1.00	----- -----	--------	1.00	----- -----	--------
2	2,029 (5.76)	33,169 (94.24)		0.85	0.80 – 0.91	<0.001	0.87	0.81 – 0.93	<0.001
3	2,015 (5.17)	36,992 (94.83)		0.76	0.71 – 0.81	<0.001	0.77	0.71 – 0.82	<0.001
4	1,992 (4.56)	41,668 (95.44)		0.67	0.63 – 0.71	<0.001	0.65	0.60 – 0.69	<0.001
5 (Most Deprived)	1,971 (3.95)	47,917 (96.05)		0.57	0.54 – 0.61	<0.001	0.54	0.50 – 0.58	<0.001
Sex			<0.001			**<0.001**			**<0.001**
Male	4,835 (4.57)	100,882 (95.43)		0.81	0.78 – 0.84	<0.001	0.81	0.78 – 0.85	<0.001
Female	5,019 (5.60)	84.651 (94.40)		1.00	----- -----	--------	1.00	----- -----	--------
Age (Years)			<0.001			**<0.001**			**<0.001**
<50	673 (16.00)	3,533 (84.00)		4.03	3.69 – 4.41	<0.001	2.61	2.36 – 2.88	<0.001
50 – 59	1,686 (10.15)	14,923 (89.85)		2.39	2.25 – 2.55	<0.001	1.74	1.63 – 1.86	<0.001
60 – 69	3,602 (7.42)	44,959 (92.58)		1.70	1.61 – 1.78	<0.001	1.46	1.39 – 1.54	<0.001
70 – 79	3,080 (4.51)	65,214 (95.49)		1.00	----- -----	--------	1.00	----- -----	--------
80 – 89	784 (1.61)	47,773 (98.39)		0.35	0.32 – 0.38	<0.001	0.47	0.43 – 0.51	<0.001
90+	29 (0.32)	9,131 (99.68)		0.07	0.05 – 0.10	<0.001	0.12	0.08 – 0.18	<0.001
Diagnosis Year			<0.001			**<0.001**			**<0.001**
2012	861 (2.68)	31,269 (97.32)		0.24	0.22 – 0.26	<0.001	0.21	0.19 – 0.22	<0.001
2013	991 (3.06)	31,437 (96.94)		0.28	0.26 – 0.30	<0.001	0.25	0.23 – 0.27	<0.001
2014	1,129 (3.47)	31,432 (96.53)		0.32	0.29 – 0.34	<0.001	0.28	0.26 – 0.30	<0.001
2015	1,439 (4.40)	31,230 (95.60)		0.40	0.38 – 0.43	<0.001	0.36	0.33 – 0.38	<0.001
2016	2,097 (6.36)	30,889 (93.64)		0.60	0.56 – 0.63	<0.001	0.55	0.52 – 0.58	<0.001
2017	3,337 (10.23)	29,276 (89.77)		1.00	----- -----	--------	1.00	----- -----	--------

Table 4.3 Continued

	Number (%) Utilising a Novel Therapy n = 9,854 (5.04)	Number (%) Not Utilising a Novel Therapy n = 185,533 (94.96)	P Value[1]	Unadjusted OR	95% CI	P Value[2]	Adjusted OR	95% CI	P Value[2]
Ethnicity			<0.001			**<0.001**			**<0.001**
White	8,812 (4.90)	170,878 (95.10)		1.00	----- -----	--------	1.00	----- -----	--------
Other Ethnic Group[4]	821 (11.72)	6,186 (88.28)		2.57	2.39 – 2.78	<0.001	1.97	1.81 – 2.15	<0.001
Missing/unknown[5]	221 (2.54)	8,469 (97.46)		0.51	0.44 – 0.58	<0.001	0.53	0.46 – 0.61	<0.001
Rural/Urban Indicator			<0.001			**<0.001**			**<0.001**
Rural Village, Hamlet & Isolated Dwellings	911 (5.92)	14,473 (94.08)		1.15	1.07 – 1.24	<0.001	0.89	0.82 – 0.96	0.005
Rural Town & Fringe	950 (5.05)	17,848 (94.95)		0.97	0.91 – 1.05	0.495	0.88	0.81 – 0.95	0.001
Urban City & Town	4,126 (4.77)	82,379 (95.23)		0.92	0.88 – 0.96	<0.001	0.88	0.84 – 0.92	<0.001
Urban Conurbation	3,867 (5.18)	70,833 (94.82)		1.00	----- -----	--------	1.00	----- -----	--------
Stage			<0.001			**<0.001**			**<0.001**
I	384 (1.18)	32,024 (98.82)		0.15	0.14 – 0.17	<0.001	0.13	0.12 – 0.15	<0.001
II	437 (2.85)	14,884 (97.15)		0.38	0.34 – 0.42	<0.001	0.38	0.34 – 0.42	<0.001
III	2,307 (6.26)	34,535 (93.74)		0.86	0.82 – 0.90	<0.001	0.92	0.88 – 0.97	0.003
IV	6,545 (7.21)	84,240 (92.79)		1.00	----- -----	--------	1.00	----- -----	--------
Unknown[6]	181 (0.90)	19,850 (99.10)		0.12	0.10 – 0.14	<0.001	0.30	0.25 – 0.34	<0.001
CCM (Between 78 to 6 Months Prior to Diagnosis)			<0.001			**<0.001**			**<0.001**
0	7,248 (6.78)	99,622 (93.22)		1.00	----- -----	--------	1.00	----- -----	--------
1-2	2,134 (3.50)	58,890 (96.50)		0.50	0.47 – 0.52	<0.001	0.72	0.68 – 0.76	<0.001
3+	472 (1.72)	27,021 (98.28)		0.24	0.22 – 0.26	<0.001	0.45	0.41 – 0.50	<0.001
Multiple Tumours			<0.001			**<0.001**			--------
1	8,340 (5.24)	150,790 (94.76)		1.00	----- -----	--------	-----	----- -----	--------
>1	1,514 (4.18)	34,743 (95.82)		0.79	0.74 – 0.83	<0.001	-----	----- -----	--------

Table 4.3 Continued

	Number (%) Utilising a Novel Therapy n = 9,854 (5.04)	Number (%) Not Utilising a Novel Therapy n = 185,533 (94.96)	P Value[1]	Unadjusted OR	95% CI	P Value[2]	Adjusted OR	95% CI	P Value[2]
Histology			<0.001			**<0.001**			**<0.001**
NSCLC NOS	1,082 (1.30)	81,871 (98.70)		0.12	0.11 – 0.12	<0.001	0.18	0.17 – 0.20	<0.001
NSCLC Adenocarcinoma	7,012 (10.15)	62,097 (89.85)		1.00	----- -----	--------	1.00	----- -----	--------
NSCLC Squamous	1,731 (4.17)	39,803 (95.83)		0.39	0.36 – 0.41	<0.001	0.47	0.44 – 0.50	<0.001
NSCLC Large Cell	29 (1.62)	1,762 (98.38)		0.15	0.10 – 0.21	<0.001	0.14	0.09 – 0.20	<0.001

[1]Chi-square P value
[2]Bolded P values are from LRT of the variable's contribution to the model. Unbolded P values are from a test of whether the OR is different from 1.
[3]For diagnosis year 2012, IMD_2010 was used and for diagnosis years 2013 - 2017, IMD_2015 was used.
[4]Other ethnicity refers to Asian/British Asian, Black/African/Caribbean/Black British, mixed/multiple ethnic groups and other ethnic groups.
[5]Missing/unknown ethnicity refers to missing and unknown ethnicity.
[6]Unknown stage refers to tumours where stage is missing and unstageable tumours.
CCM: Charlson Comorbidity Index; IMD: Index of Multiple Deprivation; LRT: Likelihood ratio test; NOS: not otherwise specified; NSCLC: non-small cell lung cancer; OR: odds ratio; 95% CI: 95% confidence interval.

4.6.3 Primary Analysis: Any Novel Anti-Cancer Therapy Restricted to Stage IV

Descriptive Statistics: 90,785 patients (46.5%) were diagnosed with a stage IV cancer - the largest proportion of any stage classification. Of these, 6,545 patients (7.2%) utilised a novel anti-cancer therapy. Utilisation of any novel anti-cancer therapy by deprivation was as follows: 9.6% (IMD 1 least deprived); 8.2% (IMD 2); 7.4% (IMD 3); 6.6% (IMD 4); and 5.6% (IMD 5 most deprived). Males were less likely to utilise novel anti-cancer therapies (6.2% of males compared to 8.4% of women). Treatment utilisation was also higher in the younger age groups, specifically those aged <70 years, with the proportion of patients receiving these treatments being highest in the <50 years age group (13.2%). Drug utilisation increased over time; 8.6% of the stage IV patients diagnosed in 2012 utilised these treatments compared to 32.8% by 2017. More patients from non-white ethnic classifications utilised these treatments (16.4% compared to 7.0% of white patients), while the uptake by rural/urban location was evenly distributed.

Multivariable Model Deprivation Associations: In the multivariable model, utilisation of any novel anti-cancer therapy reduced as the level of deprivation increased: 9.6% (IMD 1 least deprived); 8.2% IMD 2; 7.4% IMD 3; 6.6% IMD 4; and 5.6% (IMD 5 most deprived). Reduced utilisation of any novel anti-cancer therapy by deprivation status was more marked than with the whole cohort multivariable model. There was a significant association between deprivation status and novel anti-cancer therapy utilisation (LRT $p \leq 0.05$) as well a trend for reduced utilisation across all deprivation categories. The magnitude of the association in deprivation was large. Those patients residing in the most deprived areas were 55% less likely to utilise these therapies compared to those living in the least deprived areas (IMD 5 vs IMD 1 mvOR; 0.45, 95% CI 0.41, 0.49) (Table 4.4). The test for linear trend was significant ($p \leq 0.05$).

Table 4.4 Likelihood (OR and 95% CI and p values from logistic regression) of utilising any novel anti-cancer therapy (targeted therapy, immunotherapy, or biologic) deprivation and adjusted for: sex, age, diagnosis year, ethnicity, rural/urban indicator, multiple tumours, and comorbidities in stage IV patients with a NSCLC diagnosed between 01/01/2012 - 31/12/2017 (n = 90,785)

				Unadjusted			**Adjusted**		
	Number (%) Utilising a Novel Therapy n = 6,545 (7.21)	Number (%) Not Utilising a Novel Therapy n = 84,240 (92.79)	P Value[1]	OR	95% CI	P Value[2]	OR	95% CI	P Value[2]
IMD[3]			<0.001			**<0.001**			**<0.001**
1 (Least Deprived)	1,261 (9.56)	11,932 (90.44)		1.00	----- -----	--------	1.00	----- -----	--------
2	1,355 (8.17)	15,224 (91.83)		0.84	0.78 – 0.91	<0.001	0.82	0.75 – 0.89	<0.001
3	1,361 (7.38)	17,090 (92.62)		0.75	0.70 – 0.82	<0.001	0.72	0.66 – 0.78	<0.001
4	1,320 (6.57)	18,785 (93.43)		0.66	0.61 – 0.72	<0.001	0.58	0.54 – 0.64	<0.001
5 (Most Deprived)	1,248 (5.56)	21,209 (94.44)		0.56	0.51 – 0.60	<0.001	0.45	0.41 – 0.49	<0.001
Sex			<0.001			**<0.001**			**<0.001**
Male	3,112 (6.22)	46,899 (93.78)		0.72	0.69 – 0.76	<0.001	0.72	0.68 – 0.76	<0.001
Female	3,433 (8.42)	37,341 (91.58)		1.00	----- -----	--------	1.00	----- -----	--------
Age (Years)			<0.001			**<0.001**			**<0.001**
<50	503 (20.40)	1,963 (79.60)		3.78	3.39 – 4.21	<0.001	3.47	3.09 – 3.89	<0.001
50 – 59	1,162 (13.20)	7,640 (86.80)		2.24	2.08 – 2.42	<0.001	2.11	1.95 – 2.29	<0.001
60 – 69	2,395 (10.19)	21,107 (89.81)		1.67	1.57 – 1.78	<0.001	1.68	1.57 – 1.79	<0.001
70 – 79	1,952 (6.35)	28,810 (93.65)		1.00	----- -----	--------	1.00	----- -----	--------
80 – 89	509 (2.39)	20,828 (97.61)		0.36	0.33 – 0.40	<0.001	0.37	0.34 – 0.41	<0.001
90+	24 (0.61)	3,892 (99.39)		0.09	0.06 – 0.14	<0.001	0.09	0.06 – 0.13	<0.001
Diagnosis Year			<0.001			**<0.001**			**<0.001**
2012	560 (3.75)	14,392 (96.25)		0.23	0.21 – 0.26	<0.001	0.20	0.18 – 0.22	<0.001
2013	696 (4.72)	14,035 (95.28)		0.30	0.27 – 0.32	<0.001	0.26	0.24 – 0.29	<0.001
2014	755 (5.00)	14,347 (95.00)		0.31	0.29 – 0.34	<0.001	0.28	0.26 – 0.31	<0.001
2015	980 (6.31)	14,542 (93.69)		0.40	0.37 – 0.44	<0.001	0.37	0.34 – 0.40	<0.001
2016	1,407 (9.07)	14,099 (90.93)		0.60	0.56 – 0.64	<0.001	0.57	0.53 – 0.61	<0.001
2017	2,147 (14.34)	12,825 (85.66)		1.00	----- -----	--------	1.00	----- -----	--------

Table 4.4 Continued

	Number (%) Utilising a Novel Therapy n = 6,545 (7.21)	Number (%) Not Utilising a Novel Therapy n = 84,240 (92.79)	P Value[1]	Unadjusted OR	Unadjusted 95% CI	Unadjusted P Value[2]	Adjusted OR	Adjusted 95% CI	Adjusted P Value[2]
Ethnicity			<0.001			**<0.001**			**<0.001**
White	5,809 (6.96)	77,660 (93.04)		1.00	----- -----	--------	1.00	----- -----	--------
Other Ethnic Group[4]	583 (16.39)	2,975 (83.61)		2.61	2.39 – 2.87	<0.001	2.26	2.04 – 2.50	<0.001
Missing/unknown[5]	153 (4.07)	3,605 (95.93)		0.57	0.48 – 0.67	<0.001	0.45	0.38 – 0.53	<0.001
Rural/Urban Indicator			<0.001			**<0.001**			**<0.001**
Rural Village, Hamlet & Isolated Dwellings	629 (8.54)	6,773 (91.46)		1.15	1.05 – 1.26	0.003	0.92	0.83 – 1.02	0.109
Rural Town & Fringe	628 (7.06)	8,270 (92.94)		0.93	0.85 – 1.02	0.130	0.86	0.78 – 0.94	0.002
Urban City & Town	2,727 (6.73)	37,789 (93.27)		0.87	0.84 – 0.94	<0.001	0.86	0.81 – 0.91	<0.001
Urban Conurbation	2,561 (7.53)	31,448 (92.47)		1.00	----- -----	--------	1.00	----- -----	--------
Multiple Tumours			<0.001			**<0.001**			**<0.001**
1	5,726 (7.37)	71,957 (92.63)		1.00	----- -----	--------	1.00	----- -----	--------
>1	819 (6.25)	12,283 (93.75)		0.84	0.78 – 0.90	<0.001	1.15	1.07 – 1.25	<0.001
CCM (Between 78 to 6 Months Prior to Diagnosis)			<0.001			**<0.001**			**<0.001**
0	5,020 (9.13)	49,943 (90.87)		1.00	----- -----	--------	1.00	----- -----	--------
1-2	1,280 (5.01)	24,260 (94.99)		0.52	0.49 – 0.56	<0.001	0.62	0.58 – 0.66	<0.001
3+	245 (2.38)	10,037 (97.66)		0.24	0.21 – 0.28	<0.001	0.31	0.27 – 0.36	<0.001
Histology			<0.001			**<0.001**			-------
NSCLC NOS	729 (1.86)	38,542 (98.14)		0.12	0.11 – 0.13	<0.001	-----	----- -----	--------
Adenocarcinoma	5,014 (13.72)	31,527 (86.28)		1.00	----- -----	--------	-----	----- -----	--------
Squamous	787 (5.57)	13,333 (94.43)		0.37	0.34 – 0.40	<0.001	-----	----- -----	--------
Large Cell	15 (1.76)	838 (98.24)		0.11	0.07 – 0.19	<0.001	-----	----- -----	--------

[1]Chi-square P value

[2]Bolded P values are from likelihood ratio tests of the variable's contribution to the model. Unbolded P values are from a test of whether the OR is different from 1.

[3]For diagnosis year 2012, IMD_2010 was used and for diagnosis years 2013 - 2017, IMD_2015 was used.

[4]Other refers to Asian/British Asian, Black/African/Caribbean/Black British, mixed/multiple ethnic groups and other ethnic groups.

[5]Missing/unknown ethncity refers to missing and unknown ethnicity classifications. Abbreviations: CCM: Charlson Comorbidity Index; IMD: Index of Multiple Deprivation; NOS: Not otherwise specified; NSCLC: Non-small cell lung cancer; OR: Odds ratio; 95% CI: 95% Confidence interval.

4.6.4 Exploratory Analysis: Any Novel Anti-Cancer Therapy Sub-Groups

Descriptive Statistics: Of the total study population: 2.5% utilised a targeted therapy; 1.2% a targeted therapy other than erlotinib; 2.2% an EGFRi; 0.3% an ALKi; 0.5% a biological; and 2.3% an immunotherapy. When exploring utilisation of novel anti-cancer therapies sub-groups by sex, it was observed that utilisation was higher for females than males for targeted therapies, including EGFRi and ALKi (Figure 4.4a). Patterns by age generally showed that utilisation was highest in the <60 age group (Figure 4.4b) For all therapy sub-groups, utilisation was higher in the non-white population than the white population despite white patients constituting most of the dataset (Figure 4.4c). Further information on demographic and clinical characteristics by therapy sub-groups is provided in Appendix 4.10.

a)

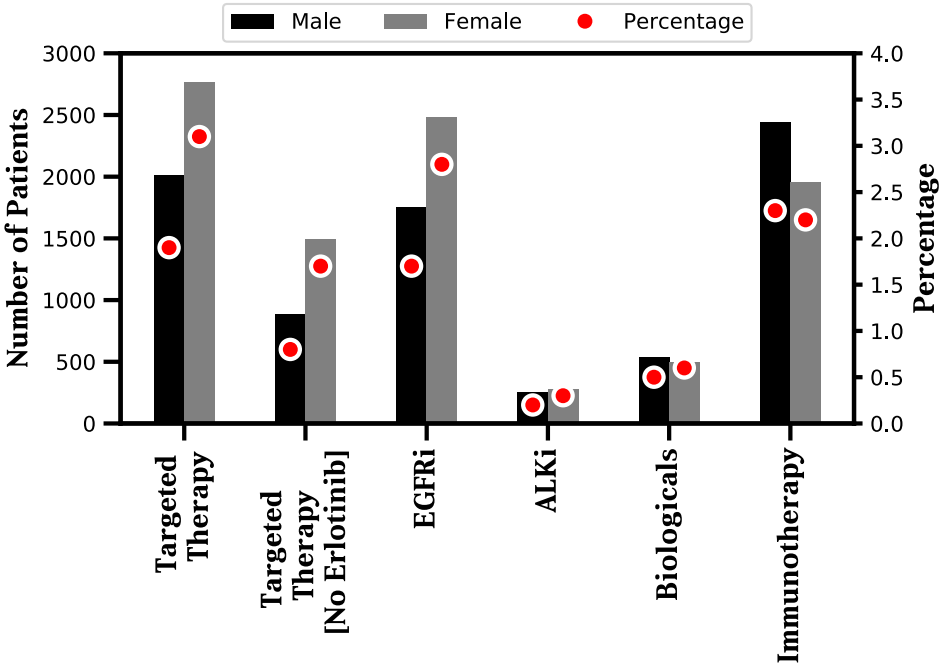
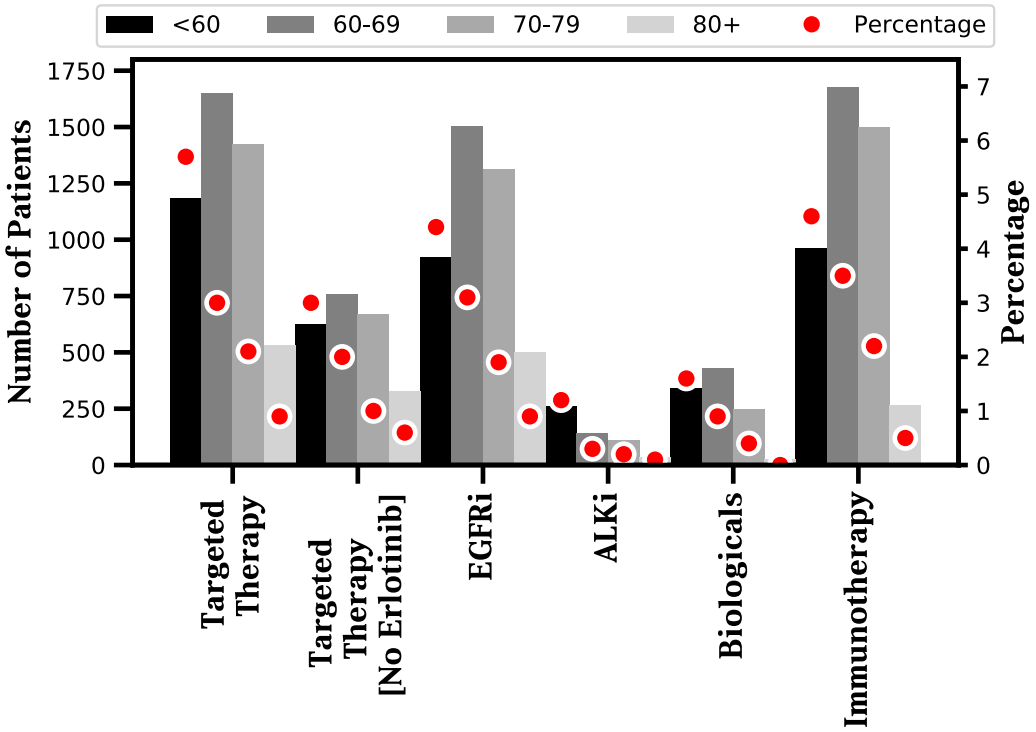


Figure 4.4 Novel anti-cancer therapy sub-classification utilisation by **a)** sex

b)



c)

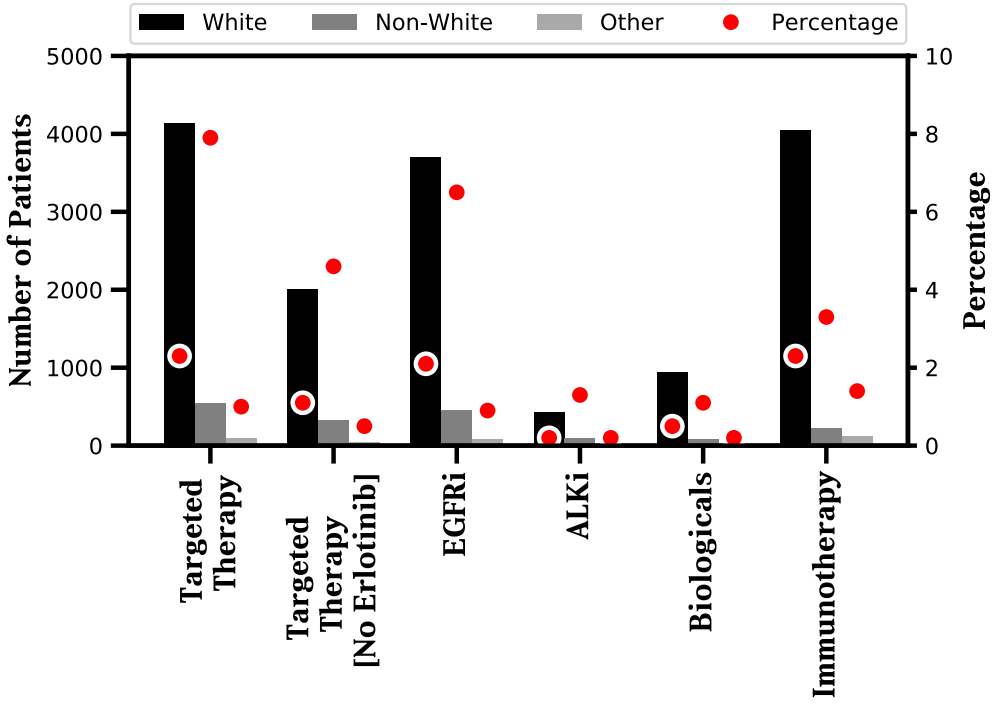


Figure 4.4 Novel anti-cancer therapy sub classification utilisation by **b)** age; and **c)** ethnicity.

Multivariable Model Deprivation Associations: When exploring the minimally adjusted models for novel anti-cancer therapy sub-groups, patterns of reduced treatment utilisation with increased deprivation were seen across all therapy sub-groups. There was a trend for decreasing utilisation with increasing deprivation was observed across all deprivation categories in all novel anti-cancer sub-groups. The association of deprivation and novel anti-cancer therapy utilisation was significant overall for all therapy sub-groups (LRT $p \leq 0.05$). However, association of the IMD 2 deprivation category and whether the adjustedOR (adjOR) was different from 1 was not significant in ALKi (LRT $p = 0.138$) and biologic (LRT $p = 0.854$) sub-groups. Patients residing in the most deprived areas were 60% less likely to utilise any targeted treatment compared to those residents in the least deprived areas (IMD 5 vs IMD 1 adjOR; 0.40, 95% CI 0.36, 0.44). When exploring targeted therapy sub-classification by drug class, patients residing in the most deprived locations compared to the least deprived were: 58% less likely to utilise EGFRi (IMD 5 vs IMD 1; adjOR 0.42, 95% CI 0.38, 0.46); 68% less likely to utilise an EGFRi that was not erlotinib (IMD 5 vs IMD 1; adjOR 0.32, 95% CI 0.27, 0.37); and 71% less likely to utilise an ALKi (IMD 5 vs IMD 1; adjOR 0.29, 95% CI 0.22, 0.38). Reduced utilisation for the most deprived patients was also observed in biological (IMD 5 vs IMD 1; adjOR 0.43, 95% CI 0.35, 0.53) and immunotherapy (IMD 5 vs IMD 1; adjOR 0.58, 95% CI 0.53, 0.64) treatments. For all therapy sub-groups, the pattern of reduced treatment utilisation with increasing deprivation was seen across all deprivation categories. Also, tests for linear trend in all sub-group models were significant ($P \leq 0.05$) (Figure 4.5, Appendix 4.11).

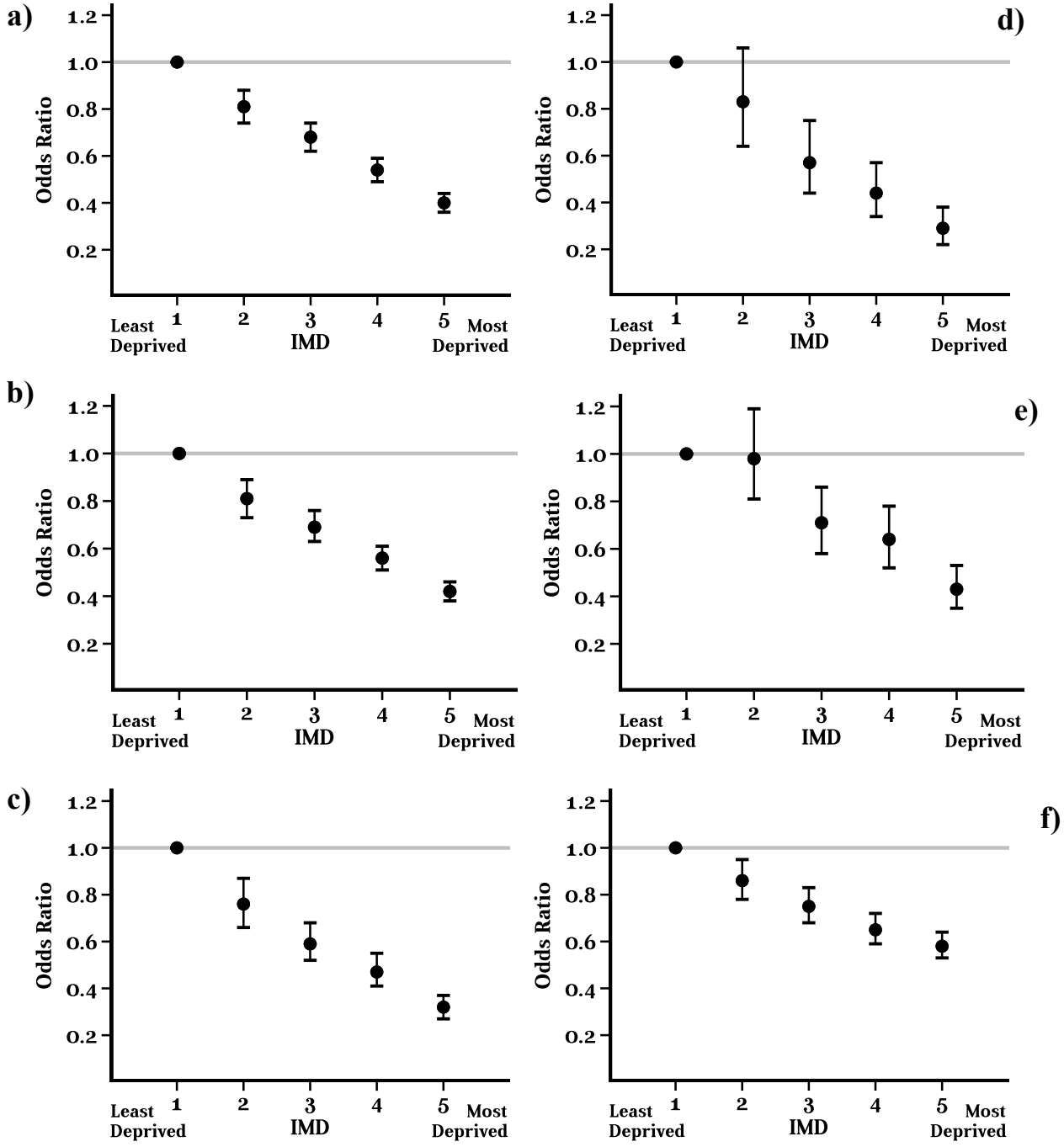


Figure 4.5 Any novel anti-cancer therapy sub-group utilisation by deprivation (IMD 1 least deprived; IMD 5 most deprived) for **a)** Any targeted therapy; **b)** EGFRi **c)** EGFRi without erlotinib; **d)** ALKi; **e)** Biologicals and f) Immunotherapy. All models were adjusted for: sex, age (<60, 60-69, 70-79, 80+), and ethnicity apart from models e) biologicals and f) immunotherapy which were only adjusted for age and ethnicity. IMD: Index of multiple deprivation.

4.7 Discussion

4.7.1 Main Findings

This population-based cohort study of 195,387 patients reports the associations between deprivation and novel anti-cancer therapy utilisation in NSCLC patients diagnosed in England between 01/01/2012 and 31/12/2017. Significant and consistent socio-economic inequalities in treatment utilisation were observed with a trend of decreasing odds of utilisation with increasing deprivation. When all patients were considered, those resident in the most deprived areas were 46% less likely to utilise any novel anti-cancer therapy than those resident in the least deprived areas. This pattern in utilisation persisted when restricting the cohort to stage IV patients (the group in which these treatments are most often used). Inequalities in treatment utilisation were observed across all sub classifications of the novel treatments (any targeted therapy - EGFRi or ALKi, biological, or immunotherapy). Inequalities were similar with targeted treatments (60% less likely to receive) and biologicals (57% less likely to receive) but were smaller in immunotherapy utilisation (42% less likely to receive). However, the small numbers of patients who received these therapies mean these results should be interpreted with a degree of caution in terms of the precision of these estimates.

These findings show associations that are much stronger and more convincing than those conclusions from Chapter 2 which reported that there are socio-economic inequalities in lung cancer treatment utilisation. The findings presented in this Chapter expand on this work and demonstrate even within the context of the UK NHS where treatment is free at the point of delivery, deprivation appears to be an important factor when considering NSCLC novel treatment utilisation. These results are important given the increasing focus towards the personalisation of care in NSCLC and the growing number of targeted treatments available. Attention is thus needed to address these inequalities so that, in time, all patients - regardless of socio-economic background can benefit from their use.

The proportion of the cohort utilising a novel anti-cancer therapy overall was small (n = 9,854; 5.0%). When compared to the international lung data reported in Chapter 2 and despite the different time periods and patient cohorts investigated, this study found: lower overall biological utilisation (0.5%) versus international data ranging from 3.1% - 34.7%; lower targeted therapy utilisation (2.5%) versus international data ranging from 9.2% - 24%; and comparable immunotherapy utilisation (2.3%) versus international data (2.3%). As the studies detailed in Chapter 2 were broad and differed in their size, methods, and denominator population (it is assumed that studies included patients where biomarker testing, where feasible,

was also carried out prior to testing access, though these details of treatment receipt with testing data has rarely been reported to date), hence it is not possible to compare this study directly to others. However overall, there was a consistent theme that utilisation of novel anti-cancer therapies in NSCLC is low, which fits with the rarity of some tumour biomarker potential.

4.7.2. Interpretation of Findings

These findings indicate that despite significant improvements in lung cancer treatment and prognosis, in no small part due to the development of novel drug therapies, socio-economic inequalities still prevail. From this work, the finding that there are socio-economic inequalities in novel anti-cancer therapy utilisation fits with the previous literature documenting large socio-economic treatment inequalities in both conventional and more novel lung cancer treatments such as SABR (221,469). Additionally, the findings adds weight to the challenges of NSCLC treatment inequalities, which, in theory could help explain inequalities in survival outcomes (151). Without knowing the optimal level of utilisation within this population, it is unclear whether the promise of precision medicine to improve outcomes has materialised. For those patients in receipt of a novel anti-cancer therapy, this may be the case. However, it is possible for example, that utilisation in the least deprived population may be higher than is shown in this analysis. Additionally, it is also possible that many more patients who would be able to benefit are not at present receiving treatment. The findings here further suggest that for those patients residing in deprived areas, there are additional barriers which may further compromise the promise of this new era of treatment in this patient population.

When exploring possible explanations for these findings, it can be helpful to consider that utilising a targeted NSCLC treatment has three steps (each with its own potential for inequality generation): (i) undertaking a predictive biomarker test; (ii) having a targetable tumour identified by a predictive biomarker test; and (iii) receiving a novel treatment (Figure 4.6). The influence of each step in these findings is now discussed.

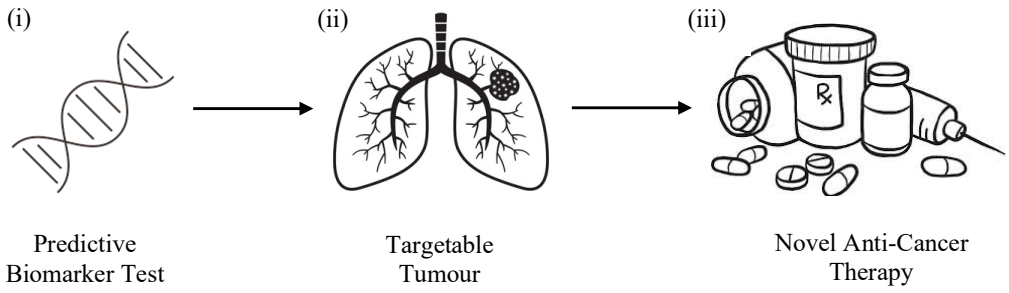


Figure 4.6 Steps to utilising a NSCLC novel anti-cancer therapy

A targetable tumour forms the basis of the precision medicine promise - that is, having a genetic predisposition towards a targeted tumour mutation.

This study did not explore the role of step (i) predictive biomarker testing as information on these tests is not recorded on national databases. It is therefore not possible to discern the importance of access to testing as a potential treatment barrier. However, previous work has identified inequalities in biomarker use. For example, Illei *et al.* (2018) detailed how older, male smokers were less likely to receive ALK testing (343), and Chapter 2 of this thesis found evidence for reduced utilisation across of several predictive biomarkers in patients with low SES (Appendix 2.1) (231). Additionally, CRUK concluded that a suspected 3,500 patients who should have been eligible for targeted treatment in 2014 did not receive these owing to 24,000 biomarker tests not being undertaken (based on estimated demand) (344). As it is already known that there are country specific differences in predictive testing availability in Europe (470), it is possible that testing referral may be socio-economically patterned. Further investigation of this is warranted, especially given that biomarker testing in the future is going to become more important and complex (e.g. whilst biomarkers such as EGFR may become standard practice, this is not necessarily true of newer biomarkers and where testing pathways require services outside the NHS, inequality generation is possible). The role of molecular tumour boards (470) and MDT discussions may prove informative in determining equitable treatment access.

Assuming patients have accessed a predictive biomarker test, this leaves the potential for two other steps in treatment initiation where socio-economic inequalities could occur. The first of these is the presence of a targetable tumour. Considering precision medicine logic, this step is arguably most influential for stratifying patients to determine treatment success. There are certain patient factors associated with an increased likelihood of tumour mutation and thus, such population groups would be predicted more likely to benefit from targeted treatment. Hence it is reasonable to hypothesise that there would be increased treatment utilisation in such patient demographics. For example, young females of Asian ethnicity are more likely to have an EGFR mutation and EGFRi susceptibility (196,471). The influence of patient characteristics (amongst other clinical characteristics) was observed in this study. Patients who were female, younger, non-white ethnicity and who had an adenocarcinoma histology and/or no comorbidities were more likely to utilise novel anti-cancer therapies. Whilst the results support the role of biological factors in determining treatment utilisation, they cannot account for *all* the inequalities observed. For example, when biological factors were included as confounders

in the multivariate models, associations between SES and novel anti-cancer treatment utilisation remained and in fact, were essentially unchanged. Biological factors alone therefore do not fully explain the inequalities seen.

The hypothesis that a targetable biology is not the only treatment determinant is further supported when examining the result for utilisation of immunotherapy when compared to utilisation of a precision medicines. Immunotherapy (specifically ICIs use) in cancer is not known to have biological drivers (unlike with precision medicines). Despite these biological differences, SES inequalities in treatment utilisation were observed in both treatment groups (though the magnitude of the association in immunotherapy was somewhat less marked than with precision medicines). Hence, when biological factors are taken out of the treatment equation, inequalities persisted. This provides further evidence that SES - or something that it is a marker for - is also important in determining treatment utilisation; having a targetable tumour biology is not the only treatment barrier.

Smoking is associated with some NSCLC sub-types (472) and in some instances, immunotherapy success (473). Smoking is also a risk factor for NSCLC diagnosis (105) and a variable known to be SES linked (106). Consequently, many NSCLC studies examine smoking given its potential as a confounding factor to explain associations between SES and treatment utilisation. Whilst these analyses could not directly measure smoking or include it as a confounder for treatment utilisation in multivariable models, it is still possible to somewhat explore smoking effects on SES utilisation through the analysis that were undertaken. For example, comparison of EGFRi utilisation (EGFR mutations are not smoking related) (471) when compared to immunotherapy utilisation (some evidence of increased efficacy in smokers) (473) in this analysis showed that both treatments were subject to inequalities by deprivation. Additionally, when comparing squamous cell (smoking related) to adenocarcinomas (less strongly smoking linked) (472), SES inequalities were observed regardless. It appears therefore that smoking variations in NSCLC sub-types, as with different tumour biology or novel anti-cancer therapy sub-type does not offer an alternative explanation to account for SES inequalities observed.

Finally, socio-economic inequalities could occur at the point of treatment receipt. Exploration of a range of novel anti-cancer therapies in this analysis was useful for examining this treatment barrier as regardless of therapy sub-type, socio-economic inequalities were observed. The consistency of the findings of SES inequalities indicates that this is systemic issue in NSCLC.

The influence of deprivation was found to be stronger when erlotinib was removed from targeted therapy analyses. This may be explained by the fact that when first licenced, erlotinib was approved regardless of EGFR status (BR-21 trial demonstrated longer survival for patients treated with erlotinib and failed to out rule this effect in EGFR negative patients) (353). It was only later that licensing changes would recommend EGFR testing prior to use (353). Hence early SACT data likely contains non-selective erlotinib use and in the analyses here, a less strong association was seen between SES and erlotinib utilisation than for some other drugs.

It is not clear why inequalities in lung cancer novel anti-cancer treatment remain so stark. Healthcare system factors, including the role of the professional may be important. There have been numerous changes in NSCLC NICE guidance in the last decade owing to the ever-growing number of novel anti-cancer therapies being licensed. Biomarker testing has too become more common place but molecular test availability is known to be variable (especially across Europe) (470). This has resulted in a situation where decision making is not always clear and requires the clinician to be abreast of changes within the field. This may mean that the potential for prescribing bias and stereotyping based on characteristics that are socio-economically patterned (e.g. fitness to undergo treatment), even in the presence of a clinical biomarker result, is still therefore possible. Additionally, patient involvement in decision making and their attitudes towards novel anti-cancer therapies may be significant. Some studies have already shown that treatment refusal in lung cancer can have links to a low SES (as defined by education status) (474), hence further investigation is needed. The role of the healthcare system and patient factors are discussed further in Chapter 5 where comparisons between cancer types are also made.

Also, of note to this analysis was the counterintuitive finding that novel anti-cancer therapy utilisation was observed in stage I (n= 384) and stage II (n=437) NSCLC, even though during the time frame of this analysis, such treatments were not licenced in these patient populations (Appendix 4.1). It is impossible to know why these 821 patients with early stage disease utilised a novel anti-cancer therapy. Most likely this reflects the fact that the registry records stage at diagnosis, rather than stage of disease at treatment. Hence many of these therapies may have been used in a metastatic setting following disease progression. Additionally, it is feasible that these numbers may capture some off licence compassionate novel anti-cancer therapy access use. On the other hand, this data could reflect errors with SACT data recording (either with regards to cancer staging, novel anti-cancer therapy use – or both). Overall, the proportion of early stage patients utilising novel anti-cancer therapies was low and similar associations with

SES were observed when the dataset was restricted to stage IV disease only, hence the impact of inclusion of these 821 patients on the results is likely small.

4.7.3 Strengths and Limitations

To date, this is the largest analysis of NSCLC novel anti-cancer therapy use in the English publicly funded NHS and the second largest study internationally after Verma *et al.* (2019) (303). The population-based nature of the dataset minimises selection bias and the large size provides the potential to explore specific groups of therapies.

It should be acknowledged, however that this study has several methodological limitations, which relate to the use of observational design and quality of the data source (SACT and NCRD data), and these have already been discussed (see Chapter 3, Section 3.7.3). A number of these pertinent limitations are discussed again briefly here.

Firstly, it is theoretically possible that the results are artefacts of data collection and therefore do not accurately reflect real-world English treatment utilisation practice. On balance, this conclusion seems highly unlikely. The magnitude of the findings, their statistical significance, along with their persistence in a range of restricted analyses of different sub populations, disputes the fact that the findings occurred by chance.

Second, the measure of SES has limitations. This study used deprivation was used as a marker of SES - and only the income domain of the IMD was used. Hence there are elements of SES, beyond that captured by the variable of deprivation that this study was not able to capture. IMD was also measured at the area not patient level so may not accurately represent SES at the level of the individual - the ecological fallacy is thus a risk. Whilst IMD is a valid measure of SES, its limitations when taken together, means the indices may misclassify the SES of certain groups/individuals (475) thereby leading to the introduction of random error in these analyses. This limitation was previously highlighted in Chapters 2 and 3 and is discussed further in Chapter 5.

Third, there may be some misclassification of a “NSCLC”. This arises because there were many patients (n=82,953; 42.46%) who were classified as having a NOS lung cancer. Had the specific histology of these patients been known, it is possible that a proportion of these patients would have been found to have not had a NSCLC. If this was the case, the associations described here are likely an underestimation of the true frequency of novel anti-cancer therapy utilisation in a

NSCLC population. A specific example of where this may have occurred is with the coding of the 8000 (general malignant neoplasms) histology code in the NSCLC definition in this analysis. The 8000 code lacks specificity and likely includes patients who never had their lung cancer histologically confirmed - usually older people and those diagnosed at a later stage. This means that the analyses may include some patients who did not have a NSCLC, hence there may be an underestimation of overall utilisation within a NSCLC population. However, it seems unlikely that this would bias results as there is no evidence that such a histological classification is socio-economically patterned. The decision to retain the 8000 code within the analysis was taken because firstly, it was deemed likely that a significant number of these cases would have in fact have been a NSCLC had they had histologically been confirmed. And secondly, removing it would result in loss of substantial number of patients (n = 38,411; 19.7%).

Fourth, there may be some misclassification of both adenocarcinomas and squamous cell histology in sensitivity analyses. This is because non-specific codes have historically been used (8010 carcinoma NOS code, and later from 2001 onwards, the 8046 NSCLC code) as opposed to adenocarcinoma and squamous type NSCLC codes (461). It is therefore possible that the adenocarcinoma sensitivity analysis does not include all adenocarcinomas in the dataset as those coded as "NSCLC NOS" will have been missed. This means that there may be an over estimation of therapy utilisation in the adenocarcinoma population. Additionally, as the non-squamous sensitivity analysis included the "NSCLC NOS" code, it is possible that some squamous tumours are present. Hence there may be an underestimation of overall utilisation within the refined non-squamous sensitivity results. The decision to run sensitivity analyses in this way was taken as it was not possible to deduce the level of detail require from non-specific codes. It is unlikely that such an approach will have introduced bias into the results; non-specific versus specific histology coding is not likely to be socio-economically patterned.

Fifth, it is possible that analyses included therapies not used for the primary invasive lung cancer diagnosis (e.g. bevacizumab may have been used for a previously diagnosed ovarian cancer rather than the lung cancer included in the dataset). However, applying time limits to treatment receipt should have limited this likelihood and hence occurrences are likely to be small and not socio-economically patterned. There is also the possibility that some utilisation data may have been missed for therapies which are primarily used at a later stage in treatment pathways (second/third line or beyond) due to the limits placed by the year post diagnosis time frame in this analysis. However, as second occurrences of a primary cancer are not that common, this was not considered problematic.

It is likely that the factors considered in this analysis do not fully account for the wider determinants which may also play a role in novel anti-cancer treatment utilisation. A clear example of this is the comorbidity data. Underreporting of comorbidities by the CCM in the lung data appears to be particularly stark as 106,870 (54.7%) patients were reported to have no comorbidities. It is likely that there is residual confounding from an unmeasured component of comorbidities in this analysis. Additionally, this analysis does not consider another "fitness for treatment" measure - performance status. This data field was requested from ODR but was supplied from the SACT not NCRD dataset (see Appendix 3.2 for the data dictionary at the time of the data request). This meant that performance status was reported for patients only at the start of a SACT treatment cycle. As over 80% of patient did not have a SACT record, missing data with this variable was high (31.9% of patients with a SACT record). This meant that it was not possible to include performance status in multivariable models. As performance status is another important consideration for assessing eligibility for treatment and is considered in the licence for novel anti-cancer therapies of interest in this analysis, lack of inclusion of this treatment determinant is a limitation of this analysis. It is therefore possible that this analysis does not fully account for the consideration of all factors in treatment utilisation. However, given the size of the SES association seen, it is unlikely that comorbidities and performance status alone could entirely explain these associations. Furthermore, the multivariable model did control for the other factors that are likely to be correlated with "fitness for treatment" (e.g. age).

Additionally, this analysis did not consider further factors which may also influence associations between deprivation and treatment utilisation such as predictive biomarker testing and/or results, smoking status, and patient preference regarding treatment choices. Influences on smoking in treatment utilisation both by tumour biology and patterning with deprivation have already been previously described (472). Smoking status, whilst now available for access from NHS Digital (NLCA dataset), was not available at the time of the ODR application and this was confirmed with PHE colleagues back in early 2019 (see Appendix 3.2 for the data dictionary extract at the time of data request). Likewise, no predictive NSCLC biomarker testing data was available for release in early 2019, only data on breast cancer biomarkers (HER2, ER, and PR) (see Appendix 3.2). In the time that has since elapsed, NSCLC biomarker testing data in England has been reported as part of a spotlight audit (193) – so testing data may now be available. NHS Digital's new dataset, the NDRS Somatic Molecular Testing dataset (476) provides details on molecular testing data collected directly from molecular diagnostic

laboratories in England. Future data requests with linkage to this new dataset could (depending on the completeness of the testing data) now enable exploration of novel anti-cancer therapy utilisation with/without prior biomarker testing access moving forward. Finally, as the SACT dataset lacks reporting of data on patient factors around declining a treatment even if it was offered, it is therefore not possible to assess the role of shared decision making between patients and clinicians which may also be important in determining utilisation - and this also limits the certainty of outcomes. Combined, all these factors mean that the analysis is likely affected by uncontrolled confounding.

Missing data was evident especially with regards to stage. This is a known problem in registry data but does reflect the fact that not all patients are histopathologically staged. However, as such data are unlikely to be missing at random, it is considered inappropriate to drop this information from the dataset as this will introduce bias. Instead, steps were taken to explicitly highlight when data was missing (often as a separate "missing" category) and this information was used in the model analyses. SACT issues with data coverage prior to 2014 have been discussed in detail previously (see Chapter 3, Section 3.7.3) but are also relevant here as a further limitation. Additionally, low rates of patients with a SACT data record (n=50,044) may have impacted the overall findings. Whilst those patients without a SACT record were not "missing" as such (there may have been a valid reason for why there was no drug record for that patient (e.g. patient not offered systemic drug therapy or patient choice to decline treatment)), an absence of a drug record does mean however that it is not possible to determine if the low overall utilisation of novel anti-cancer therapies observed in this analysis reflects actual low utilisation overall or low recording of utilisation. Had such data been available, associations with utilisation of novel anti-cancer therapies *may* have been attenuated and overall rates of utilisation may have been found to be higher. Comparisons were made of the demographics and clinical characteristics of those patients with and without a SACT record (Appendix 4.5) and similarities were noted. This means that it is difficult to hypothesise why under reporting of novel anti-cancers therapies would be associated with SES in so much that this has introduced a false association into the multivariable models. As SACT recording may improve moving forward as the scheme becomes longer established, it will be important to repeat these analyses for more recent years of data.

The small patient numbers from non-white ethnicities are a further limitation of the data. Ethnicity data in English cancer registries historically has not been well recorded. Whilst completeness of specific data items has improved in time, there is still the potential for

misclassification - especially of non-white ethnicities (379,477). This is a drawback given that ethnicity may be important in certain tumour biomarker susceptibility. For example, Black African and East Asian ethnicities sub-groups are relevant to lung studies such as this one given that the Black African ethnicity has no mutation associations whilst East Asian ethnicity can be linked to an increase in EGFR mutation. A *post-hoc* analysis of any targeted therapy utilisation was explored in these ethnic groups was performed and utilisation found to be similar, though slightly higher in the Asian population (Black n = 129 (7.0%); Asian n = 310 (9.9%)). However, the small number of patients in each of these ethnic groups recorded in the dataset means that such comparison should be interpreted with caution.

Finally, it is important to highlight that treating NSCLC is a rapidly evolving field. Diagnosis and treatment data reported in the study are now five years old and may not accurately reflect current practice. The data here provides a snapshot of prescribing practices at the time of the study.

Despite these limitations, the data presented in this Chapter provides important real-world information on the utilisation of novel anti-cancer therapies in a publicly funded healthcare setting - such prescribing practices are not captured in randomised controlled trials which typically include highly selected patient populations (478).

4.8 Conclusions

This retrospective analysis of English cancer registry and linked SACT data has shown that deprivation is an important factor in NSCLC novel treatment utilisation, even in a publicly funded healthcare system. Given that NSCLC treatment is rapidly evolving, and guidelines are quickly changing, further work - exploring *why* such significant treatment inequalities persist is needed to ensure the promise of stratified NSCLC treatment to improving outcomes for all is fully realised. The goal of precision medicine should be that all patients who stand to benefit from novel anti-cancer therapies should receive them - regardless of socio-economic position in society.

Chapter 5. Discussion & Conclusions

5.1 Introduction

The previous Chapter concluded the empirical work of this thesis with a second cohort study of the NCRD and SACT data with a NSCLC focus. It highlighted, that in comparison to the findings of the HER2+ breast cancer cohort study (Chapter 3), these socio-economic inequalities in the utilisation of novel anti-cancer therapy in NSCLC are much more pronounced. This Chapter synthesises the entirety of the work presented in this thesis and discusses the findings in relation to the initial research question proposed back in Chapter 1.

Aim: To discuss and evaluate the thesis results.

Objectives:

1) Situate research findings in the context of the original thesis aim and previous literature.
2) Discuss the relevance and significance of the research findings.
3) Explain the research findings within the context of the thesis's strengths and limitations.
4) Comment on the implications of this work for clinical practice and policy.
5) Provide further research ideas for ongoing work in this field.

The Chapter begins with a brief review of the main findings before bringing these results together with a discussion of their meaning and significance when considered in the context of the overall thesis aim. A discussion of the thesis strengths and limitations follows. Actions resulting from this work are then suggested. This includes a discussion of unanswered questions, implications for both policy and practice as well as future research directions. The Chapter ends with concluding remarks.

5.2 Summary of Main Findings

This thesis sought to establish whether (and if so - to what extent) there are socio-economic inequalities in the utilisation of novel anti-cancer therapies and their associated predictive biomarkers.

This aim was met through three studies which are summarised below.

5.2.1 Systematic Review & Meta-analysis

Chapter 2 addressed the first two objectives of this thesis: to systematically review the existing evidence exploring possible associations between SES: and (i) predictive biomarker testing utilisation; and (ii) novel anti-cancer therapy utilisation.

The systematic review and meta-analysis found that in studies published to December 2019:

- Low SES (measured with a variety of markers) was associated with modestly lower predictive biomarker test utilisation (14%; 95% CI 0.71, 1.05) and significantly lower novel anti-cancer therapy utilisation (17%; 95% CI 25%, 9%) than high SES.
- SES associations with treatment utilisation varied by cancer type. A low SES (when compared to high SES) was associated with significantly lower novel anti-cancer therapy utilisation in lung cancer (29%; 95% CI 0.51, 1.00) and a modestly lower treatment utilisation (7%; 95% CI 0.78, 1.10) in breast cancer.
- Published studies mostly used datasets from non-publicly funded healthcare systems in the USA.

5.2.2 Trastuzumab Utilisation in HER2+ Breast Cancer

Chapter 3 addressed the third objective: to determine the association of SES (measured in terms of the IMD income domain of area of residence at diagnosis) on trastuzumab and conventional treatment (breast cancer directed surgery and chemotherapy) utilisation for a HER2+ breast cancer population in England.

The population-based cohort study found that:

- Women residing in the most deprived areas were 8% (95% CI 15%, 1%) less likely to utilise trastuzumab and 21% (95% CI 27%, 14%) less likely to utilise breast cancer directed surgery compared to women residing in the least deprived areas.
- Patterns of chemotherapy utilisation by SES were complex and varied by patient age, and whether the woman had undergone breast cancer surgery.

5.2.3 NSCLC Novel Anti-Cancer Therapy Utilisation

Chapter 4 addressed the final thesis objective: to determine the association of SES (measured in terms of the IMD income domain of area of residence at diagnosis) on novel anti-cancer therapy utilisation for a NSCLC population in England.

This population-based cohort study found that:

- Patients residing in the most deprived areas were 46% (95% CI 50%, 42%) less likely to utilise any novel anti-cancer therapy compared to patients residing in the least deprived areas.
- Inequalities by deprivation of a similar magnitude were also observed when restricting to a stage IV cohort.
- In exploratory analyses, associations with deprivation were stronger with targeted therapy compared to immunotherapy utilisation.

5.2.4 Overall Findings

Overall, the work presented in this thesis found:

- A lower SES was associated with a reduced likelihood of utilisation of both novel anti-cancer therapy and their associated predictive biomarkers.
- A reduced likelihood of treatment utilisation with a lower SES existed regardless of the nature of the healthcare system.
- A reduced likelihood of treatment utilisation with a lower SES was observed across a range of solid tumour cancers - though magnitude varied by cancer type.
- As with conventional cancer treatments and despite treatment advances, it appears that lower SES is a potential barrier to fair treatment utilisation.

5.3 Interpretation of Findings

This thesis has built on work which has previously described reduced conventional cancer treatment utilisation in low SES groups (92,216–218,220,221,223). It also adds to the wider of body of work describing longstanding health inequalities across the UK (12,34,113–115,117) as well as health inequalities in cancer (121,122). This thesis found that in England, reduced treatment utilisation with a low SES was observed, despite the presence of a publicly funded healthcare system where care is based on need, not an individual's ability to afford treatment. These findings suggest that should inequalities be eradicated (i.e. rate of treatment utilisation in the least deprived patients is applied to all other IMD quintiles), a further 41 patients would have been anticipated to receive trastuzumab in HER2+ breast cancer (total population of 40,179 patients) and a further 3,205 patients would have utilised a novel anti-cancer therapy in a NSCLC context (total population of 195,387 patients).

The observational drug utilisation data reported in this thesis is subject to two types of error: bias and uncontrolled confounding; moreover, results could be due to chance. However, it seems unlikely that these factors could provide an alternative explanation for the results obtained. This is because, where possible methods deployed minimised these risks. For example, it is hard to hypothesise a systematic bias whereby treatment is inaccurately recorded, and the level of inaccuracy is associated with SES. Furthermore, multivariable analyses conducted in Chapters 3 and 4 minimise the risk of uncontrolled confounding as potential confounders were considered which may also impact the reduced utilisation observation with novel anti-cancer therapies. However, it was not possible to control for all potential confounders (e.g. frailty, smoking, and performance status) either not measured by NCRD and SACT data (Appendix 3.2) or not provided in usable format due to substantial levels of missingness. These variable/factors may be found to attenuate results had they have been included in the analyses, but this was deemed unlikely given the strength of associations seen in some analyses and the fact that other measures (e.g. age and comorbidities), were taken into consideration where applicable. Finally, the fact that this thesis used population-based data and reports large sample sizes throughout (Chapter 1 included data on over a million patients, with Chapter 3 and 4 providing additional information on over 40,000 and 195,000 patients respectively), it seems reasonable to conclude that the likelihood that associations recorded here are due to chance are small. Rather, the finding that a lower SES is associated with reduced treatment utilisation is indeed real.

The findings from this work raise the question as to why a low SES increased the likelihood of reduced treatment utilisation. As already detailed in Chapter 1, the influence of SES is likely multifactorial given that good health, and access to healthcare (including treatment) has several determinants (e.g. biology, personal characteristics, physical environment, and the social and economic environment). These determinants may also be associated with SES. An example of such an association is that a later cancer stage presentation has been previously linked to low SES - and this will have implications for whether treatment choice is wide and/or curative (332). To account for this consideration, where possible this thesis has explored the role that cancer staging on novel anti-cancer therapy utilisation (see Chapter 4 stage IV NSCLC analysis), concluding that socio-economic gradients in utilisation do occur in later stage groups. Patient factors such as poor health literacy (e.g. lacking the capacity to recognise cancer symptoms and an inability to communicate with clinicians) may also increase the likelihood of a later stage diagnosis and limit how a patient can navigates the healthcare system - including decisions around novel anti-cancer therapy acceptance or rejection (131,474). Furthermore, if more

biologically more aggressive tumour phenotypes are present in those with low SES – for example through: environmental exposure (e.g. working conditions) (479), pollution (480), and increased smoking levels (e.g. TP53 mutations) (481) - then treatment options in such patient groups may also be reduced. If stage and grade data – which are routinely recorded in cancer registry datasets do not fully capture tumour aggressiveness, this explanation will be hard to test at present with real-world datasets. Multimorbidity or a reduced "fitness for treatment" provides a further SES barrier (lifestyle factors e.g. poor diet, smoking, drinking alcohol, poor working conditions are linked to a low SES) (482).

The environment may determine the ease at which patients can access care. In publicly funded healthcare systems where costs are not an issue, distance to services, transport access, childcare, and flexible work conditions may still result in socio-economic barriers for low SES patients and this may, in turn limit treatment utilisation (131). Geographical constraints may be particularly restrictive to patients of low SES trying to access to specialised or centralised services, not present in their local vicinity. This is problematic as greater distance has been shown to be associated with less treatment (224). Increased travel distances and associated travel costs (as per Figure 1.3 in Chapter 1) can therefore exacerbate existing treatment disparities between patients of different socio-economic groups as patients of lower SES may be constrained by suboptimal or overstretched service provision in their local vicinity (483). Additionally, the local environment and primary healthcare availability can result in diagnosis delays, later staging, and limited treatment options, especially in deprived communities (484). The role of timely lung cancer diagnosis provides an example given that patients often repeatedly visit their GP before referral or investigation is made (485). Delays in primary care referral for a suspected lung cancer chest X-ray may therefore be significant on both staging and onward treatment and are likely SES linked given region disparity in care provision (though it is worth noting that the analyses in this thesis were adjusted for stage). Variation in chest X-ray rates have been observed (486) and may form an additional barrier for patients from lower SES backgrounds if such referral is not a priority in their local vicinity. Furthermore, the impact of delayed presentation can be large, especially if cancers are diagnosed through an emergency admission as this is associated with poorer outcomes (including with treatment options given the need for urgent management) (487). Patients from a low SES have been shown to be at increased risk of emergency presentation (487) and this can impact on treatment options. Finally, implicit societal bias and stereotyping based on SES (along with other factors such as age and comorbidity) could influence treatment decisions which may also, in part, account for reduced treatment utilisation in low SES groups (214,227,488).

The influence of SES in reducing treatment utilisation is likely additive, intersectional (31,59) and cumulative over the life course. Combined, the numerous interaction of SES with overall health, symptom recognition, navigation of the healthcare system amongst other factors outlined already, may explain why inequalities in novel anti-cancer therapy utilisation are observed despite the promise of precision medicine to provide a solution (26,38). An example of the intersectionality of the role of SES across the life course could be that children and adolescents of low SES report lower health status and more risk behaviours which by adulthood may have translated into premature chronic morbidity. These inequalities may result in individuals becoming patients in later life - with in an increased likelihood of cancer diagnosis, including at a later stage. Hence treatment choices (including novel therapies) may be more limited and health literacy surrounding informed treatment decision may be lacking to make an informed decision (26).

The role of individual socio-economic barriers to treatment can be complex. One example of this is with the interaction between education and income. There may be situations where a patient does understand the consequences of inaction, whether this be a failure to attend screening or follow up appointments or to turn down an expensive treatment when offered. In such instance, patients may have adequate health literacy. However, financial constraints - both direct (e.g. drug costs) and indirect costs (e.g. travel costs, supportive care, lost work, income, and savings) (489) – lead to situations where financial matters are entwined with the patient's clinical decision making processes (490). Such scenarios also present dilemmas for clinicians when confronted with patients unable to afford their medical care (490). It is therefore important to differentiate between health literacy and financial health literacy when discussing treatment access as an individual's ability to make sound healthcare decisions is also based on access to resources available to them (491). Patients facing financial toxicity (492) may have to make tough compromises such as between paying for food or a hospital bus fare. A cancer diagnosis alone has been shown to have a financial impact (83% of UK cancer patients experience some sort of financial impact post diagnosis) (493) – without the added external financial constraints of a current cost of living crisis (494). In times of financial uncertainty, it is not difficult to understand why patients, of varying levels of deprivation, may make different clinical decisions given material and psychological stressors (492,495). It is therefore important that the wider political and economic determinants are also considered, beyond individual SES, as further barriers to equitable novel anti-cancer treatment utilisation.

Overall, the findings suggest that the overarching promise that precision medicine will improve outcomes for all, curtail health inequalities and improve public health (172,225,226) has not yet materialised. Whilst in theory, precision medicine with its simplification of treatment access to a clinical decision alone (the presence or absence of a predictive biomarker) removes the potential for inequality in utilisation - this is not always observed in practice. This is likely due to the fact, that as outlined in Chapter 1, genetics are only one component of a myriad of factors which interact to influence health and healthcare access (26,57,183). It may be that in the UK, where there is an element of freedom in exercising clinical judgement, implicit societal biases e.g. by age (and by extension SES) may help explain the existence of inequalities even when guidelines are prescriptive and predictive biomarker results clear. It can be helpful to conceptualise precision medicine as a "downstream" intervention serving to "fix" inequalities at the point of drug utilisation but doing so fails to consider the "upstream" fundamental causes of inequality generation in the first instance (27), i.e. the inequitable access to social resources (e.g. wealth, income, and education which ultimately helps individuals to seek treatment). These fundamental causes have been argued to drive existing health inequalities and perpetuate the ICL (availability of good medical care varies inversely with the need of the population served) repeating in numerous settings - including in a precision medicine treatment era (55,57).

Precision medicines can be considered an example of an IGI (a new medical advance which has the potential to widen inequalities – see Chapter 1) (56). Whilst this thesis did not find any evidence of treatment inequality widening in novel anti-cancer therapy use (when compared to conventional cancer therapy utilisation) it did note the overall persistence of inequalities in treatment utilisation based on lower SES. However, it was also noted that associations of lower SES on treatment utilisation varied by cancer type. This suggests that the extent to which precision medicine presents as an IGI and the potential of the precision medicine revolution to impact on treatment utilisation may vary on a cancer-by-cancer basis. In HER2+ breast cancer, the novel anti-cancer therapies utilised by all patients may be considered more successful. Whilst this thesis did find inequalities in trastuzumab utilisation, the extent of these associations with deprivation were modest when compared to conventional surgical utilisation associations (Chapter 3) and especially when compared to the finding of the NSCLC analysis (Chapter 4). There may be different factors at play in determining treatment access between drug and surgical treatments as previously discussed in Chapter 3. Different factors may also be at play between different cancers (e.g. breast cancer incidence is generally associated with affluence and lung cancer, with deprivation). It is however feasible that in cancers in which novel anti-cancer therapy prescribing and predictive biomarker testing is well established and clinical

judgment is guideline driven, the temporary inequality caused by the introduction of an intervention (as described in the ICL) (55) over time reduces. This may be because treatment access widens so that all socio-economic groups can benefit from the change in standard clinical practice (see the inverse equity hypothesis, Chapter 1) (230). Breast cancer may provide such an example of this occurring given that hormone testing was well established before HER2 testing was added to NICE guidance.

In contrast, for cancers where treatment is evolving rapidly and treatment guidelines are constantly changing (e.g. NSCLC), it may still be some time before patient stratification based on tumour biomarkers mitigates the risk of lower treatment utilisation associated with lower SES (26). The results here show a far greater level of inequality between those of high and low SES in this lung cancer. NSCLC has witnessed numerous NICE guideline changes in recent years, as well as new biomarker identification and therapies targeting these being developed. This pace of change marks somewhat of a contrast to HER2+ breast cancer (where HER2+ testing was introduced in the late nineties and determined for all early breast cancer patients from 2005 onwards (496); trastuzumab was licenced by NICE in 2002 for advanced disease (68)) and this may in part explain differences in the magnitude of the inequalities observed in these cancers. It therefore may be some time before a reduction of inequalities is observed in NSCLC treatment - if at all. Such a hypothesis still needs consideration in the wider determinants of health - the factors other than SES which complicate treatment utilisation, and which may not be comparable between cancer types.

5.4 Significance of Findings

The findings of this thesis further understanding of the real-world use of novel anti-cancer therapies and the testing of the associated biomarkers. This thesis thus reports novel findings. It has detailed the first systematic review and meta-analysis to examine whether there are socio-economic inequalities in novel anti-cancer therapies and their associated testing. Additionally, it reports the first of two population-based studies exploring socio-economic inequalities in novel anti-cancer treatment utilisation in an England in a publicly funded healthcare setting. Outputs from this thesis can be broadly considered important for several reasons as outlined below:

Significance for Patient/Patient Population: The presence of unequal precision medicine utilisation raises the question as to what the consequences are for patients if they do not receive a novel anti-cancer treatment. Given that novel anti-cancer therapies have been documented to

provide patients with opportunity for improved quality of life (e.g. less toxic side effects) and improved survival outcomes (2,240,241), differential treatment receipt on the basis of SES may result in differential patient outcomes. Specifically, survival and quality of life may be shorter for both specific individual patients and on average across the patient population.

Significance for the NHS & Wider Society: It has already been hypothesised that inequalities in treatment may result in inequalities in outcomes (e.g. mortality and survival) (92,126,127,143,213). It is plausible that unfair novel anti-cancer treatment utilisation may further explain, in part, English cancer survival and mortality inequalities. For example, it is already known that the UK has lower cancer survival rates than many other similar countries (497,498), and the English analyses may also help explain this fact along with why, even though survival is improving over time, the disparity between the UK and other countries is not narrowing. Additionally, given the increasing implementation of pharmacogenomics over a range of medical conditions, learning how to manage and prevent socio-economic inequality development with precision medicine implementation in areas other than oncology will also be important.

Significance for Further Research: Knowledge that current precision medicine treatment provision, regardless of medical advance or healthcare system, appears to be unfair, raises the question as to what will happen if no changes are actioned - and not just in breast and lung cancers. Given the pace of change in the precision medicine field (new biomarkers and therapies are continually being discovered) (235) action to address these treatment inequalities is urgently required. Additionally, the presence of SES inequalities in predictive biomarker testing highlights an additional barrier to equal treatment utilisation which has not been raised before in previous work on treatment inequalities focusing on conventional cancer treatments. Consideration of the role of biomarker testing as an additional barrier to treatment utilisation will be important in future work moving forward. Suggestions for the specifics of future work are discussed in more detail later in this Chapter (see Section 5.8.2).

5.5 Strengths

The methodology presented in this thesis has many strengths, and these have already been described in the relevant Chapters (2 - 4). This section therefore focuses on the overall strengths of the thesis.

The large-scale coverage of real-world utilisation of novel anti-cancer therapy utilisation is a strength of the work. The population-based nature of the findings provides not only a comprehensive understanding of current novel therapy utilisation in clinical practice but also some confidence that the associations observed are likely to be real and not merely the result of bias, uncontrolled confounding or chance as previously discussed. Conducting observational analyses as part of drug utilisation research studies goes beyond smaller, more selective data obtained from randomised controlled trials, or single centre case series, which often lack representativeness and external validity (354). Furthermore, addressing the research question for the first time within England as a whole (Chapters 3 & 4), as well as at a global scale (Chapter 2) is a particular strength. This is because it provided exposure to a range of contexts in which novel anti-cancer therapies are used - and in which different factors, may or may not, have had the potential to influence the outcome. This approach also enabled the evaluation of potential associations across a range of healthcare systems (publicly and non-publicly funded), something that may be particularly pertinent given novel therapy utilisation and the cost of novel anti-cancer treatments. The thesis findings therefore have local, national, and international relevance.

The use of a relatively new, "big dataset" to answer this research question is a further strength. Traditionally pharmacoepidemiology research has tended to focus on primary care health data, where most drugs are prescribed (e.g. pharmacy dispensing information, GP records, administrative claims data) (499). In contrast, this thesis demonstrates how drug utilisation studies can benefit from accessing routine in-patient health services records on treatments used in hospitals (like SACT) which, when linked with established cancer registries, increases the information available and offers opportunity to improve understanding of cancer treatment in a real-world setting. The work also indicates how the SACT database could be further applied to answer other pertinent research questions about expensive novel therapies, not used in primary care, can be studied.

Combined these strengths have resulted in outputs (see Publications & Conference Presentations from the thesis, Page v) which lay the basis for further work exploring how to guide public policy during this new precision treatment era. Example suggestions are discussed further in Section 5.7.

5.6 Limitations

Despite the numerous strengths of this work, this thesis is also subject to several limitations. These have been discussed where appropriate in each analysis Chapter (2 - 4). The limitations provided below thus reference the overall methodology used for the thesis.

The main concern when undertaking drug utilisation research is whether the data are the best fit to answer the research question in hand. In terms of Chapter 1, setting out to synthesise previously reported data are always subject to the fact that there is no guarantee that such data will provide opportunities for comparisons if both the number of studies along with the lack of comparability of individual studies does not allow for this. Whilst there was no documented evidence of between-study publication bias in Chapter 2 (Appendix 2.13), there was a possibility that within-study reporting bias was present (though this was not measured directly as statistical methods to determine this are not well developed) (500) and that this could have influenced the direction of the outcomes. Additionally, it is not possible to directly compare the systematic review finding with the English data analyses due to inconsistency in the way SES is recorded. However, in common with the searches included in the systematic review, NCRD and SACT are limited by the data fields recorded at any given time and these fields may not be consistently reported from record to record. For example, whilst the thesis aimed to explore socio-economic inequalities this was only possible through an income lens because only the income domain of IMD is routinely available in the cancer registry. This has consequences for inferences made from these results - see Section 5.8.2 for further discussion.

NCRD and the linked SACT dataset provide a large, comprehensive, population-based database which undergoes rigorous quality assurance measures. Despite these advantages, data quality is far from perfect (as for any real-world routine health dataset). A clear example of this, described as a limitation in both Chapter 3 and 4, was the reporting of comorbidities. Population-based cancer registries generally do not collect information on the presence and severity of coexisting medical conditions despite the potential of existing medical conditions to impact cancer treatment options (501). Rather, the measure of comorbidity (of which there are numerous examples e.g. Charlson, Elixhauser) is derived from linkage with other sources of health data (see Lüchtenborg *et al*. for an example of this approach (502)). In the NCRD, the Charlson Comorbidity Index (CCM) is used and is derived from HES (379). Whilst the CCM (437) is a valid measure of patient comorbidities, its measurement is known to be crude (the weighted measure only pertains to only 16 comorbidities occurring during in-patient hospital care) (104,438). Therefore, under ascertainment of comorbidities is likely to be common among

cancer patients recorded on the NCRD; patients never admitted to hospital for one of the conditions included in the CCM, or who were admitted but not during the time window HES records were scrutinised, or who had another condition which is not included on the CCM but that impacts cancer treatment decisions, will have their true level of comorbidities under-ascertained. Hence, even if patients are treated for a comorbidity in primary care, their CCM score can be zero (223).

Limited availability of comorbidity data in this thesis means that it is likely data in both Chapters 3 and 4 is subject to residual confounding by comorbidities. The influence of "fitness to treatment" which arguably may be less of an issue with novel anti-cancer therapies than with conventional cancer treatments, was not possible to measure in the population-based analyses undertaken. This issue has been previously detailed as problematic in other cancer data analyses (503). It is further complicated by the fact that SES maybe associated with comorbidity among cancer patients (generally those with a lower SES have poorer health status and a higher risk for serious comorbidity e.g. cardiovascular disease, chronic obstructive pulmonary disease (COPD) and diabetes) (504). A further issue arises from the fact that comorbidity can refer to both physical and mental illnesses, yet as with many measures of multimorbidity (505), the CCM lacks consideration of mental health comorbidity. As cancer diagnoses and treatment have known emotional and psychological burdens (506), the level of control a patient feels they have, emotional support system along with their management of any feelings of blame, stigma and shame from the diagnosis may all impact healthcare service and treatment engagement (507). Treatment refusal has been linked a low SES (474) and yet with current comorbidity data capture and under reporting, it is not possible to tease out the impact of this on treatment utilisation. Efforts were taken to reduce the impact of this measure on the results (especially in the absence of performance status and frailty measures) by including the CCM into multivariable models along with other factors such as "age" which can be surrogate measures for treatment fitness, however it is likely that some unquantifiable effects of comorbidity on the association between SES and treatment receipt remain. Given the overall strength of SES associations observed in Chapter 4's lung data, it is unlikely that uncontrolled residual confounding by comorbidities could entirely explain the association, though it may be an important consideration in Chapter 3's breast data where more modest SES associations were described. Given that the breast Chapter's findings were consistent with other international analyses with different patterns of uncontrolled confounding, this may be unlikely.

There are also further limitations concerning the NCRD. The first arises from the fact that this is a tumour, not patient, registry and this complicates the analysis when exploring a patient, not tumour level characteristic (e.g. SES). Analyses had to be conducted on the data at the level of the patient, not the tumour as many statistical approaches assume the unit of analysis/observations are independent. Cleaning the data to the level of the patient however incurred a loss of information on additional primary invasive tumours occurring within the same patient. For example, in the HER2+ breast cohort, 1,056 tumours from the full 263,293 breast tumours obtained from ODR were "lost" via this cleaning process. Although, there is no reason to assume that any of these additional primary invasive tumours would be biased by SES to be treated differently to the first primary tumour, it is possible that patterns of utilisation in patients with multiple tumours might be different or that the first tumour may influence treatment of the second.

As SACT is a relatively new dataset, data quality and consistency issues were anticipated. Data input is by hospital staff, often in busy clinics when a patient is present and this may mean that the data recorded is not always accurate and complete (354). For example, it has already been highlighted that for some Trusts, treatments are documented in the SACT data after a patient is known to have died. Recording issues with SACT data when treatments maybe delayed and/or omitted is problematic (401). A major concern with the SACT data, already highlighted in Chapters 3 and 4, is that prior to April 2014 data reporting for all Trusts was not mandatory (354). This means that early SACT data likely under ascertains novel anti-cancer therapy prescribing before this date. This under ascertainment is likely less problematic beyond April 2014 given that the introduction of electronic prescribing systems helped assist recording and reporting of data for national SACT uploads but this integration of electronic systems varied between individual hospital Trusts (401). However, given the cost of novel anti-cancer therapies when compared to conventional SACT treatments, it is anticipated that any under recording may be less pronounced for these treatments. Sensitivity analyses by period of diagnosis/treatment were used to account for this limitation, and it is worth noting that SES associations were found to persist regardless of the time frame considered.

Chapter 2 raised important issues about the potential for predictive biomarker testing to be an additional barrier to fair treatment utilisation in a precision medicine era. However, it was not possible to extend this focus into the later empirical work undertaken for Chapters 3 and 4 for two reasons. First, for the lung analysis (Chapter 4), biomarker data are not currently collected and/or reported in the NCRD or SACT dataset to enable analysis (correct as of the date of ODR

data application). This meant that it was not feasible in this analysis to explore utilisation of predictive biomarker tests in this cancer. Second, whilst the NCRD does report breast cancer biomarkers (ER, PR, and HER2), and these could be used as inference for test utilisation, due to limitations of time, no such analysis was undertaken. This is a limitation of this work. Additionally, it was not possible to judge whether a lack of biomarker testing was in fact the reason why socio-economic treatment inequalities were observed. Though, given that inequalities also occurred in treatments which do not necessarily require a positive biomarker status prior to drug access (e.g. some NSCLC immunotherapies and biologicals), it seems unlikely that a lack of access to testing is the root cause of all treatment inequalities reported in this thesis [Although, if the finding was due to significant inequalities in biomarker testing this would be important and as issue of concern for the NHS and patients]. What these results do highlight is that an analysis of treatment utilisation by SES stratified by a positive biomarker result would be useful to confirm if socio-economic gradients in treatment utilisation occur regardless of testing status. At the time the work was undertaken for this thesis, an analysis with biomarker data was not feasible as such data was not available for release from ODR (Appendix 3.2). Now, as the NDRS Somatic Molecular Dataset (476) has been added for linkage to the cancer registry for release by NHS Digital, an opportunity to conduct this work is possible. This will enable a timely and important review given the evolving role of biomarker testing, especially in a NSCLC context where EGFR testing is becoming standard practice, though this is not necessarily the case for newer biomarkers entering clinical guidelines.

There was a significant proportion of patients in the analytical cohorts who lacked a SACT record. It is not known if SACT records were missing for novel anti-cancer therapies or if these patients never received a SACT treatment. There is no reason to believe that a patient would or would not have a SACT record in England based on SES. The characteristics of the patients with and without a SACT record were checked and no significant differences were observed. It was therefore deemed unlikely that excluding from the analysis patients without a SACT record would introduce bias in so much that the analytical population would no longer reflect the target population the study aimed to apply its finding too. It is also a limitation of this thesis that the SACT data represents a snapshot in time. Hence all analyses reported here are time specific.

Measurement of SES is a further potential limitation. Defining an individual's SES is complex as it is hard to measure accurately and reliably, hence the SES measures described in this thesis are inferences only. Ideally, this work would have included as many SES measurements as possible, acknowledging that single measures are not interchangeable and different measures

may point towards different potential explanations for the findings. However, this was not always feasible or practical. For example, in Chapter 2, when synthesising findings from previous publications there was no level of control over the measures of SES reported (or the level at which it was assessed - primarily area-based). Furthermore, in Chapters 3 and 4, the only measure of SES available within NCRD for release was area-based deprivation (and specifically as measured using the income domain of the IMD only). A lack of individual level measures of SES is a common limitation of drug utilisation research as living in a deprived area is not the same as being deprived (508). However, individual SES is not straightforward to measure as it is a social construct which changes over time (509) and as the data are often sensitive in nature, even when asked about their SES (e.g. household income), patients may decline or exaggerate answers (510). Given the lack of individual SES data available in population-based datasets, the use of area-based measurements was expected, though this also means that SES misclassification and misinterpretation is a risk (the ecological fallacy) (511). For example, IMD is a measure of area-based level disadvantage (and in the case of Chapters 3 and 4 by income only). It therefore cannot be assumed to accurately reflect an individual person's level of deprivation as areas categorised as deprived may contain large numbers of people who are not in fact deprived and vice versa (475). Whilst a useful measure, it is also important to remember that the concordance between IMD with individual deprivation status is still not well known (512).

A small number of studies have explored the concordance of ecological to individual measures of SES outside of England and have concluded that the ecological fallacy is indeed present and agreement between area and individual-based measures is low (513,514). In England, the use of area-based measures is still common (e.g. in the NHS long-term plan), though there is some emerging evidence on the extent to which such area-based English measures reflect individual circumstance (512). For example, Ingleby *et al.* (2020) (512) assessed the concordance of area-based SES measures (including IMD) in England and Wales across six major cancers. Their work concluded that area-based income derivation was a poor predictor of individual income status, though area-based occupation and to a lesser extent, education, have a slightly higher concordance with individual based measured (512). These conclusions have important consequences on studies which examine deprivation as it is likely the area-level variable does not fully represent socio-economic variation and this can result in misleading health and deprivation inferences (512). For this analysis, their findings suggest that true associations of socio-economic measures at the individual level are likely masked.

Despite these limitations, area-based measures are widely used and worthy of their own study as a measure of interest in exploring relationships with deprivation. Even though avoiding ecological bias with aggregate measures is challenging, indices such as IMD still remain a useful tool for improving understanding of the contextual factors in inequality generation, as well as providing some ability to discern socio-economic inequalities in health databases that lack individual SES measures at present due to privacy. Many countries, including Wales therefore justify the use of the IMD as a SES proxy due to the absence of suitable individual-level SES replacement data (475). Moving forward, it is likely that exploration both individual and area-based measures will be useful for understanding and reducing socio-economic treatment variation (512).

Finally, the generalisability of these analyses to cancers other than breast and lung is uncertain, though there may be some application to other publicly funded healthcare systems.

5.7 Implications for Policy & Practice

When considering the research findings in relation to clinical practice, there are several practical applications. There is an urgent need to transition from description of inequalities towards implementing interventions which might avoid their development in the first instance (53). In relation to the types of novel anti-cancer therapies investigated here, it is acknowledged that changes in both treatment options (i.e. active treatments for cancer which, in the past, may have been managed by best supportive care) and approaches to determine who is eligible for treatment (i.e. biomarker testing) will require commitment and time for the health system to adapt to this new paradigm of treatment (515). In the first instance, from a practice perspective, policymakers and clinicians need to be aware of the potential barriers to biological and precision medicine therapy beyond patients' tumour molecular profiles and this can be started through training of clinicians in the SDoH (229). Revising guidelines to help remove some of the subjectivity in clinical decision making and to also include a focus on reducing inequalities could help assist such a prioritisation.

Secondly, if multiple barriers occur to novel therapy utilisation (including now at an additional step of predictive biomarker test utilisation), then sophisticated solutions are likely required to prevent cancer inequalities widening further. Interventions to improve access for all are likely to involve a range of levels (patient, provider, healthcare system, and society). Example solutions are suggested in Table 5.1.

There is no consensus as to which intervention is most likely to improve the socio-economic treatment gradient. Whilst all interventions have merit, those which will evoke the most impact are likely upstream interventions tackling the structural and fundamental causes (the SDoH) (27,30). These interventions need to occur long before a patient has ever received a cancer diagnosis or has a need to access a novel anti-cancer therapy. Fundamental cause interventions are likely to be most successful as they address the "why" of the problem, as opposed to trying to "fix" it after it arises. However, as societal, and political change to enable such success will take time, current downstream interventions focused on local treatment receipt improvement are still valid and arguably more significant to patients currently accessing these treatments.

Finally, changing practice is unlikely to be straightforward or simple. It is likely that inequalities and the solutions need to tackle these will need to be both local and national based. Solutions may vary based on cancer type and wider societal issues such as the healthcare system. Interventions will thus require tailoring as there is no simple one, universal solution. Overcoming traditional barriers (e.g. health literacy following a social gradient (516) and the impact this can have on a patient's ability to self-advocate and/or to seek out alternative advocacy on their behalf (517)) may compromise treatment utilisation in all socio-economic groups. Patient advocacy groups (e.g. NSCLC EGFR Resisters (518); NSCLC ALK Positive UK (519); Breast Cancer Now (520)) could assist with ensuring that novel anti-cancer therapy information is patient accessible.

At a Trust level, SACT data has previously been used to promote good practice and to highlight those failing to meet targets (e.g. companion report to the 30 day mortality data post SACT admission work) (521). Whilst target driven cancer care practice aligned to the NHS Long Term Plan objectives is now well established (120,522), there is still potential for expansion beyond a focus on referral, diagnosis, and time to treatment to also incorporating targets which consider the type of treatments received and the effect of socio-economic prescribing. There may even be scope for league table Trust level data on socio-economic associations with drug prescribing (should further work exploring Trust level inequality in SACT data be conducted) in the hopes that "naming and shaming" individual providers raises awareness of treatment inequality and provides incentives for actions for change. Such an approach has previously been used to try to stimulate improvement in cancer patient experiences e.g. Macmillan Report on league tables for the 2011-2012 National Cancer Patient Experience Survey (523). In contrast, clinical audit methods would provide a more targeted approach for structuring feedback and performance against a set of standards for delivering high quality care rather than "naming and shaming" in

national headlines (524). However, such audit approaches rely on the professionals to be motivated to change their practice and evidence on the effectiveness of such methods to date has been found to be relatively small to moderate, with the effects on patients even less clear (525). Despite this, audit and feedback is still a widely used strategy in Europe and one that may provide the impetus for policy makers to undertake a similar strategy at least on a local scale (524). Implementation of audit may therefore be best viewed as a compliment to other strategies, deployed on a case-by-case basis to tackle the extent of inequalities at hand.

Table 5.1 Possible solutions with examples of how inequalities in novel anti-cancer therapy access may be approached

Intervention Level	Solution(s)	Real-world Examples where Available*
Patient	Increased health insurance coverage in non-publicly funded healthcare systems to mitigate the cost of expensive novel anti-cancer therapies (53).	USA Medicare Expansion (Affordable Care Act) has been associated with increased cervical, breast and colorectal cancer screening in low-income adults owing to test cost reduction (co-pay or doctor visit costs) (526).
	Low-income assistance programmes for uninsured or low-income patients who are unable to afford medicines (238).	PAPs offered by pharmaceutical manufacturers and Patient Support Foundations provide eligible patients with access to brand-name outpatient chemotherapy at little or no cost has been shown to increase compliance and access at one USA public hospital in patients who otherwise would not have been able to afford treatment (527).
	Increased access to GPs to help address the ICL (55).	Changes to NHS resource allocation between 2001 and 2011 to increase NHS resources for more deprived areas in England found that geographical inequalities in mortality from causes amenable to healthcare declines in absolute terms in more deprived areas (528).
	Improved health literacy to assist patient motivation and capacity to understand and participate in shared treatment decision making (53).	Polio vaccine pamphlet written in short, simple sentences (aimed at 4th grade reading level) was preferred by low income patients in a sample in the USA (531).
	Tailor treatment information based on patient SES (e.g. beliefs and internal health locus of control) to help inform patient choice (529).	Community based, literacy sensitive and culturally tailored lifestyle intervention on weight loss and diabetes risk among a low income, Spanish speaking Latinos at increased diabetes risk found weight reduction and a drop in HbA1c compared to the control group (532).
	Evaluate financial barriers to treatment access (530).	Sheffield's Children's Hospital is using a local poverty data algorithm to identify children (and families) at risk and therefore less likely to attend for care (e.g. medical appointments). Targeted support includes free transport to the hospital and/or appointments being undertaken at the school (533).
Provider	Increased education about the potential for novel anti-cancer therapy inequalities which stem from biases, cultural beliefs, stereotypes, and learned behaviours (53).	E.g. Implementation of community service learning /SDoH to improve medical and dental student training towards people marginalised in society were found to promote a greater appreciation of the vulnerabilities faced by such population groups (534).
	Clinician training to keep up to date in terms of treatments for specific cancers (53).	
	Improve access to tertiary services for disadvantaged groups unable to travel e.g. e-consultations to reduce the need for face-to-face visits (535).	Install virtual communications for those with digital exclusions in places that people can access in their local vicinity e.g. Citizens Advice Southampton has installed video-based support in library booths to improve access to support for vulnerable, digitally excluded people (536).

Table 5.1 Continued

Intervention Level	Solution(s)	Real-world Examples where Available*
Healthcare System	Streamlined referrals (53).	In home educational intervention by lay health workers to increased adherence amongst low income, inner city, African American women to breast and cervical cancer screening found that clinical breast exams and mammograms were improved (538).
	Use of patient navigation interventions for people living in areas of disadvantage to assist with widening treatment access (537).	Telephone support interventions from prevention care managers in eleven community and migrant health centres in New York, USA improved cancer screening rates amongst minority and low income women (539).
	Monitoring/surveillance of drug usage could identify locations with underuse and or unfair service provision. This could prompt targeted investigation into barriers in these locations (238).	Implementation of clinical audits such as those carried out by the NLCA (394) which has been used to drive improvements in the quality of care for lung cancer patients e.g. making recommendations to the National Screening Committee to review screening data.
Society	Global collaboration to reduce financial barriers to drugs between countries in order to provide opportunities for increased drug distribution (540).	Price thresholds based on per capita GDP in combination with pharmacoeconomic modelling evaluations establish a value-based price e.g. (wealthier nations pay more and these higher revenues subsidise developing world drug access) e.g. hypothetical modelling for bevacizumab access in metastatic colorectal cancer in India indicates improved access under this scheme (541).
	Targeting interventions towards areas with known barriers to fair treatment access to facilitates change in socioeconomically disadvantaged populations (238).	Targeted brief behavioural counselling in general practice on low income patients on their consumption of fruit and vegetables resulted in sustained increases in dietary change consumption (542).

*Examples not necessarily derived from this topic are but provided here to show the potential expansion of solutions into this new realm of treatment. Abbreviations: GDP: Gross domestic product: HbA1C: Haemoglobin A1c; ICL: Inverse care law; PAPs: Patient Assistance Programs; SDoH: Social determinants of health: USA: United States of America

5.8 Future Research Directions

5.8.1 Unanswered Questions

The thesis has raised several questions which require further investigation. Firstly, it is important to understand why SES inequalities in novel anti-cancer therapy utilisation are present given the prescriptive nature of clinical guidelines following identification of a known tumour biology as a precursor for treatment access. Further work which investigates the wider role of the clinician, patient, and family in decision making around predictive biomarker tests and treatment utilisation may prove an enlightening starting point. This will be important given that individual choice has been shown with conventional treatment in lung cancer to be influential (474) and may also relate to SES - the same may be the case in a precision medicine context too. An example of a study exploring these wider novel anti-cancer treatment barriers could be a qualitative study using interviews and focus group methodology to understand how both patients and healthcare professionals come to make decisions regarding novel anti-cancer therapy treatment, with consideration of SES.

This thesis has focused on one area of inequalities (SES, mainly deprivation). Further work exploring inequalities in the other domains of IMD would be useful, as would consideration of other measures of SES (especially those framed at the level of the individual as opposed to the area). This would be of value, because as previously discussed, the unit of SES focused on in the English analyses does not fully explain the construct, especially at the level of the individual and this limits inferences made here. For example, if the full elements of the IMD domains were routinely available for access at ODR, it would be useful to see the influence this has on the results presented in Chapters 3 and 4. Additionally, as treatment utilisation associations were observed with other factors such as age (significant reduction in both HER2+ trastuzumab and NSCLC novel anti-cancer therapy utilisation with increasing age in the English analyses), work exploring age inequalities in receipt of these treatments is needed. Investigating the concept of ageism in novel therapy treatment utilisation would be useful as would work to consider patient views on treatment worth (e.g. whether they perceive the treatment to be worse than the cancer diagnosis). Whilst, the concept of ageism with conventional therapies is not new (543), further work developing these ideas within the context of novel anti-cancer therapies, that have a different side effect profile and treatment outcomes, also needs consideration. This could be achieved through qualitative work (e.g. interview, survey) exploring with patients of different ages who have refused treatment the rationale behind this decision. There is also the possibility to undertake some additional quantitative work using

SACT data to investigate whether dose reductions reported in SACT are age related and may therefore contribute to shared decision-making choices in such age groups.

The thesis has not explored if inequalities in utilisation results in inequalities in outcomes. It therefore remains unknown if not utilising a precision medicine may result in consequences such an reduced survival and decreased quality of life (4,544). An obvious next step would be to explore survival time following both utilisation and non-utilisation of novel anti-cancer therapies. Such an analysis is well within the remits of NCRD collection. Furthermore, as the SACT dataset is allowed to mature, in a few years' time, calculation of 5 year survival will be possible for any drug commenced after 2014.

Finally, terminology and definitions have proved somewhat of a challenge throughout this thesis especially when the terms used (e.g. targeted therapy) were not all encompassing of the nuances of precision treatments. Given the heterogenous nature of cancer, the rapid pace of change in treatments and the lack of consensus on a standardised terminology, this body of work has found that existing precision medicine definitions are not always helpful. Moving forward, a standardised, plain-language terminology would assist patients, providers, and researchers (545). However, to be successful, it would be helpful to also list individual drug and/or drug classes by standardised terminology. Given the number of treatments available and the growing number of drugs being developed, such a classification by drug name/type seems unlikely to be achievable.

5.8.2 Further Research Considerations

This section now discusses wider topics surrounding future research than those already mentioned.

SES Measures: Health (and treatment) inequalities research could benefit from a standardised SES measurement. There are arguments for use of all the three, key single measures of SES (income, employment, and education) as a new "gold standard" and these are now discussed in turn.

First, income may appear an obvious choice for further studies. This is because income is recognised as important for enabling material resource access which can in turn improve health outcomes (21,546). However, as a standardised worldwide income categorisation (i.e. what is considered "high" and "low" SES, as well as the same number of sub-groups) is unlikely,

studies instead should aim to measure income at the individual, as opposed to just the neighbourhood scale, as well as considering how income may change over time. Even if such issues were addressed, limitations such as age dependent relationships (income on average increases with age), the high non-response rate at data collection (income is often considered a sensitive discussion topic) and not addressing other assets such as health insurance or ownership of assets (wealth) would still limit its application (547).

Second, prioritisation of more employment-based measures may therefore be preferable. However, such indicators face similar drawbacks to other single measure tools. For example, classifications of "retired" or "homemaker" are hard to rank, making assessment of what constitutes "low" or "high" SES challenging (547). Additionally, for women, there is also the issue of whether SES should be assigned based on their own employment (individual) or that of a husband/partner/household (548).

The third option could be to prioritise the third element of SES - education. Indeed, this measure has been useful in discriminating different SES groups in the UK. However, the past 20 years have also seen a significant expansion of higher education, accompanied by an erosion of the concept of "graduate" jobs and careers, which has weakened the links between education and income (549). Thus, education as an SES measure may still work well for older populations, but not for younger ones.

A fourth option would be to consider an aggregate SES measure (550). However, to be effective in pooling worldwide population data, there is an assumption that datasets consistently record a range of SES factors - which is not always the case. Furthermore, with current aggregate measures, there remain challenges. For example, in the USA, aggregate measures focus on social class rather than measures of material and social deprivation which limits cross-country comparisons (547).

In the absence of finding a superior measure of SES applicable to all areas, future work could still benefit from an appreciation of the limitations of SES measures. To achieve this, researchers need to consider the SES indictors available (the constraints of the dataset access may restrict the feasibility of further measurement) and explicitly acknowledge and critically evaluate the limitations of the measure(s) used (54). This would include consideration of the following: ecological fallacy; the ability of the measure to enable cross study comparisons; the potential for differential misclassification; and issues with missing information. As all

indicators have limitations, the ability to make cross-study comparisons also requires contextualising in any future findings. Where possible, future work reporting multiple measures of SES (if available) and stratifying multivariate analyses (or testing for interactions) by other factors such as ethnic group/race, gender and age would be of value in minimising the limitation imposed by the lack of standardised "one size fits all" SES indicator (54,547). Descriptors of what each SES group pertains too, especially where sweeping classifications such as "low SES" or "high SES" are used would also be helpful.

SACT's Potential: The thesis summarises two analyses of SACT data, which whilst important, arguably only scratch the surface of how the SACT dataset could be used to investigate cancer treatment inequalities. For example, there is a wealth of data on other novel anti-cancer therapies recorded which were not explored as part of this thesis. This includes other HER2+ breast cancer treatments such as trastuzumab emtansine, pertuzumab, and neratinib. Extension to novel anti-cancer therapy use in other breast cancer sub-types would also be possible. It is anticipated that in time and with the integration of more targeted and personalised treatments into clinical guidelines, that the utility of SACT for analysing precision medicines use will grow. As treatments become more precise for rarer molecular biomarkers, data controllers will have to increasingly consider patient confidentiality of SACT data available for release.

It would be useful to perform an updated analysis (for example in 5 years' time) of both HER2+ trastuzumab and NSCLC anti-cancer therapy use. Doing so would provide time not only for improved recording of SACT data (fewer missing SACT records) but would also enable NSCLC therapies (some of which were not long licenced in this study, and which have relative low utilisation numbers at present) to become embedded into clinical practice. Larger numbers would enable expansion of the currently provided NSCLC exploratory analysis. Additionally, changes over time in drug licencing and funding may feed into drug availability - which has a role in subsequent drug utilisation. Furthermore, if the Inverse Equity Hypothesis (230) is true in novel anti-cancer treatments, an analysis in say five years' time, should in theory demonstrate lesser treatment inequalities. A repeated analysis of this type would thus enable assessment of whether the modest inequalities in breast cancer could reflect earlier trastuzumab licencing and if in time, as NSCLC also become more established, inequalities associated with lower SES in this cancer also reduce.

SACT data could also be utilised to determine treatment inequalities in cancers beyond breast and lung tumours. A second and subsequent analysis of other novel anti-cancer therapy drugs

(e.g. other anti-HER2 agents such as pertuzumab) is a likely next step. Finally, linked SACT and NCRD data could be used to conduct survival analyses following treatment receipt to examine the consequences of drug receipt/non-receipt on patient outcomes. Whilst the NCRD does document some data on quality-of-life indicators, the numbers of patients for whom this is available at present are likely too small for the measurement of this "impact" on patients to be determined.

Improving Population-based Cancer Registries: Population-based cancer registries, such as NCRD, have improved over the past few decades considering technical and scientific advances (551). Despite this, many registries have yet to realise their full research potential. Strategies for improving cancer registries further are now discussed in the context of: (i) data ascertainment; (ii) data synthesis and linkage; and (iii) and data release.

Registry data for analysis are only as good as the data input. First, allowing time in clinical practice for accurate and thorough data recording would help minimise mistakes and omission in data reporting. Investing in good data ascertainment could help improve data collection moving forward. For example, this could result in a reduction in the number of patients without a tumour stage or performance status listed or increase the number of patients with data reported in a linked SACT record – all of which were known to be limitations of this thesis' work. Second, it would also be helpful to encourage the practice of using specific codes where provided. An example of this would be when documenting lung cancer morphology and avoiding use of the 8010 carcinoma NOS code in favour of reporting the cancer instead using either the adenocarcinoma or squamous type NSCLC codes. Third, the breadth of variables available for request and subsequent linkage could be widened. It would be of benefit to have a measure of frailty along with all biomarker testing conducting. Current registry data access through NHS Digital's Data Access and Request Service does technically provide this (e.g. molecular testing data on lung cancer mutations is available in the NDRS Somatic Molecular Dataset) (476), offering scope for future analytical work. Though it has been noted that given the increasing use of precision medicine relying on biomarker status, a registry which links all relevant genomic mutations with the central cancer registry would be a "remarkable population-based research resource" (551) and should be an action point for NCRD improvement in England. Fourth, cancer registries would benefit from an additional measurement of SES. This could consist of – at a minimum – providing access to other data domains, bar income, of the IMD. Or it might include measurement of another standard for SES measurement e.g. education level. Ideally, registries should strive to have a variable of individual measurement of SES to

help minimise ecological fallacy concerns. Linkage with primary care records and a movement towards one patient record which erodes the barriers between primary and secondary care might be a starting point for the challenge of present scarcity of data on individual SES measures at a population-based scale. Fifth, improving the historic poor recording of ethnicity in the cancer registry needs prioritisation. First steps here may include where the barriers to ethnicity documentation lie and how such difficulties can be overcome (477). Prioritising the recording of ethnicity data in secondary care is required and may benefit from financial incentives to do so, as these have been proved to be successful within primary care settings under the Quality and Outcomes Framework (QOF) (552). Finally, the SACT dataset would benefit from having data ascertainment on both drug indication and line of therapy. This will help minimise drug misclassification in further analyses using drugs which can have multiple licenced indications.

The cancer registry would also benefit from an overhaul of data synthesis and linkage. As outlined in Chapter 3, the process of registry infrastructure to enable data flow is multifactorial, backdated and often requires manual upload to a secure portal or via email to NDRS. Moving forward, the registry would benefit from a simplified data extraction system, with separate data repositories which link together and do not require separate data uploads to the ENCORE platform. This would also help data to be "live" and reflect current cancer practices. [It is perhaps worth noting that as of February 2023, the most current year of cancer incidence that is available from NHS Digital is 2020, so things have in fact become worse than when data was requested for this thesis]. Overcoming the challenges of data linkage (e.g. registry responsibility, privacy, security, and quality of information) needs prioritisation, especially as most registries have been designed with a specific goal and data outside of this goal may be limited (553). Linkage of cancer data to patient records in a primary care setting, at a national level, would also be beneficial. The concept of an NHS shared summary care record (SCR) which follows a patient throughout the different organisations involved in their care and assists clinicians and professionals involved in patient care is a reality (554). However, use of the data contained within such a repository, could be improved, for example through developing a connective digital system that registries could also draw data from to aid acquisition, synthesis, and linkage.

Finally, whilst registries provide a great data repository, this needs to be accessible for release to be useful. Registries should therefore be aiming to improve access to up to date, "real time" data. Data release at the option of the patient and as opposed tumour level would also assist with analyses exploring multiple tumours within one patient.

Big Data: Further monitoring of inequalities is required and additional drug utilisation work harnessing big data will be important for obtaining such information. However, given the rapidly evolving nature of the precision oncology field, and the fact that routine datasets generally lack good data on newly licenced therapies and tests, how such data will be captured in large-scale population-based observational studies still requires some thought. Analysis of additional, new data sources, rich in biological and precision therapies or predictive biomarker tests information not recorded in SACT (e.g. England's NHS National Breast and Lung Cancer Audits along with the USA's Flatiron Health EHR database) may provide the first steps here. Although such databases also have documented limitations (e.g. the recorded level of testing and targeted therapy rates in Flatiron is lower than expected, most likely because EHR can fail to record scenarios where treatment decisions based on clinical judgement override the presence of a targetable mutation) (555). Where possible data linkage should be encouraged, especially in cancers such as NSCLC where many different types of tissue and blood-base biomarker assays (556) can lead to testing being outsourced to providers with appropriate laboratory infrastructures within the local vicinity who do not always share information in a format linked with NHS records. Insufficient lung biomarker testing data in England has already been flagged and set as a priority for the new Genomic Medicines Service's along with assessment of whether this has consequences for treatment utilisation (344). Following examples of established good practice for improved record sharing, not only for patients care but also for creating data repositories that can be utilised by researchers (e.g. Great Manchester Care Record (557) and the Great North Care Record) (558), could provide useful starting points for improving big data. Encouragement of data collection to enable audit of treatment access would inform development of solutions to respond to any inequalities noted.

5.9 Concluding Remarks

This PhD has explored the role that SES has in the utilisation of predictive biomarker tests and novel anti-cancer therapies, through a systematic review and meta-analysis as well as two observational studies. The research has highlighted several findings. First, as with conventional cancer treatment, there are socio-economic inequalities in the utilisation of both predictive biomarker tests as well as novel anti-cancer therapies. A lower SES was found to be associated with reduced treatment utilisation in all aspects of this thesis. Second, SES is an important factor in treatment utilisation even when a tumour mutation is known and implicated in treatment choice. Third, these inequalities appear to exist regardless of the nature of the healthcare system. Fourth, the magnitude of inequalities observed can vary based on cancer type - with larger socio-economic inequalities being observed in lung as opposed to breast

cancer. Moving forward in an era of increasingly biomarker driven cancer treatment, further work is needed to address these inequalities at the patient, provider, and health system level.

Chapter 6: Appendices

Appendix 2.1 Published Systematic review and meta-analysis paper for Chapter 2

Norris *et al. BMC Medicine* (2020) 18:282
https://doi.org/10.1186/s12916-020-01753-0

BMC Medicine

RESEARCH ARTICLE Open Access

Are there socio-economic inequalities in utilization of predictive biomarker tests and biological and precision therapies for cancer? A systematic review and meta-analysis

Ruth P. Norris[1,2], Rosie Dew[2], Linda Sharp[2], Alastair Greystoke[3], Stephen Rice[4], Kristina Johnell[5] and Adam Todd[1,2*]

Abstract

Background: Novel biological and precision therapies and their associated predictive biomarker tests offer opportunities for increased tumor response, reduced adverse effects, and improved survival. This systematic review determined if there are socio-economic inequalities in utilization of predictive biomarker tests and/or biological and precision cancer therapies.

Methods: MEDLINE, Embase, Scopus, CINAHL, Web of Science, PubMed, and PsycINFO were searched for peer-reviewed studies, published in English between January 1998 and December 2019. Observational studies reporting utilization data for predictive biomarker tests and/or cancer biological and precision therapies by a measure of socio-economic status (SES) were eligible. Data was extracted from eligible studies. A modified ISPOR checklist for retrospective database studies was used to assess study quality. Meta-analyses were undertaken using a random-effects model, with sub-group analyses by cancer site and drug class. Unadjusted odds ratios (ORs) and 95% confidence intervals (CIs) were computed for each study. Pooled utilization ORs for low versus high socio-economic groups were calculated for test and therapy receipt.

(Continued on next page)

* Correspondence: adam.todd@newcastle.ac.uk
[1]School of Pharmacy, Newcastle University, King George VI Building, King's Road, Newcastle-upon-Tyne NE1 7RU, UK
[2]Population Health Sciences Institute, Newcastle University Centre for Cancer, Newcastle-upon-Tyne, UK
Full list of author information is available at the end of the article

(Continued from previous page)

Results: Among 10,722 citations screened, 62 papers (58 studies; 8 test utilization studies, 37 therapy utilization studies, 3 studies on testing and therapy, 10 studies without denominator populations or which only reported mean socio-economic status) met the inclusion criteria. Studies reported on 7 cancers, 5 predictive biomarkers tests, and 11 biological and precision therapies. Thirty-eight studies (including 1,036,125 patients) were eligible for inclusion in meta-analyses. Low socio-economic status was associated with modestly lower predictive biomarker test utilization (OR 0.86, 95% CI 0.71–1.05; 10 studies) and significantly lower biological and precision therapy utilization (OR 0.83, 95% CI 0.75–0.91; 30 studies). Associations with therapy utilization were stronger in lung cancer (OR 0.71, 95% CI 0.51–1.00; 6 studies), than breast cancer (OR 0.93, 95% CI 0.78–1.10; 8 studies). The mean study quality score was 6.9/10.

Conclusions: These novel results indicate that there are socio-economic inequalities in predictive biomarker tests and biological and precision therapy utilization. This requires further investigation to prevent differences in outcomes due to inequalities in treatment with biological and precision therapies.

Keywords: Precision medicine, Molecular targeted therapy, Immunotherapy, Biological therapy, Pharmacogenomic testing, Drug utilization, Socio-economic factors, Meta-analysis

Background

Traditional cancer treatments (chemotherapy, surgery, and radiotherapy) are subject to inequalities in utilization by socio-economic status [1]. These socio-economic inequalities have persisted over time and exist across cancers, healthcare systems, and treatments [2, 3]. Individuals with a lower socio-economic status are less likely to receive conventional treatments, and this may contribute to poorer cancer outcomes in this group [4].

Increasingly, systemic treatments targeted at cancer biology (e.g., tyrosine kinase inhibitors and monoclonal antibodies) are being integrated into cancer clinical care. These agents are expensive (immunotherapy can cost, in US dollars, $100,000 per patient annually) and may only have efficacy in selected sub-populations [5]. Hence, stratifying patients by molecular pathology to predict the likelihood of tumor response and adjusting therapy accordingly is now routinely recommended (see, for example, [6]). This move towards biological and precision therapies is reflected in the cancer drug development pipeline; for example, in 2019, 450 new cancer drug candidates were immunotherapies [7].

Socio-economic inequalities in biological and precision therapy utilization remain largely unexplored. Some speculate that using molecular information to target cancer treatment potentially provides a solution to current treatment inequalities [8]. Others argue that novel cancer therapies, because of their cost, disproportionately favor those with more resources and, therefore, may widen inequalities further [9–11].

As novel cancer therapies and their associated predictive biomarker tests offer opportunities for increased tumor response, reduced adverse effects, and improved survival, it is important to understand whether there are inequalities in their receipt [12, 13]. This systematic review and meta-analysis integrated the existing research to investigate the relationship between socio-economic status and utilization of biological and precision cancer therapies and their associated predictive biomarker tests.

Methods

The review was registered with the international database of prospectively registered systematic reviews, PROSPERO (CRD42019140016), and is reported according to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) [14] (Additional file 1: Supplementary methods 1).

Search strategy and study selection

Searches were performed in seven databases (MEDLINE, Embase, Scopus, CINAHL, Web of Science, PubMed, and PsycINFO) for articles published between January 1998 and December 2019. This time period reflects the licensing and approval of trastuzumab in the USA—considered a crucial time marker in the precision therapy field. Therapies of interest included the following: targeted therapy (targeting either oncogene addiction or synthetic lethality with activity restricted to tumors with appropriate biomarker status), biologics (where no predictive biomarker is included in the license), and immune checkpoint inhibitors. Therapies targeting hormone receptors were excluded as these agents have been in use since the early 1970s [15]. Search terms covering socio-economics status, tests, and therapies were developed; a full search strategy is available in Additional file 1: Supplementary methods 2. Reference lists of eligible articles were also reviewed.

The inclusion criteria for full-text papers, published 1998 onwards and written in English, were determined as follows in terms of PICOS (Population, Intervention, Comparison, Outcome, and Setting). *Population*: solid tumor cancer diagnosis (any age or sex). *Intervention*:

receipt of either a predictive biomarker test or biological and precision therapy (or both). Studies reporting biological and precision therapies administered with an adjuvant (e.g., chemotherapy) were eligible as long as it was clear how many patients received the biological or precision therapy. Only predictive biomarker tests of pharmacological response to targeted treatment were included. *Comparison*: it was not a requirement that a comparator was reported but where noted, and the following comparator details were extracted—a clinical alternative, no biological and precision therapy and/or predictive biomarker tests, or no treatment. *Outcome*: utilization data reported by a socio-economic status measure (e.g., percent of persons living below the poverty line, median household income). *Setting*: retrospective or prospective observational design (including randomized controlled trials analyzed as observational cohorts). Full inclusion criteria are listed in Additional file 1: Supplementary methods 3.

Screening of titles and abstracts was conducted by one author (RN) only. All articles selected for full-text review were independently checked by a second author (AT). Disagreements were discussed and, if necessary, resolved with a third author (LS). Agreement between reviewers was excellent (κ = 0.93) [16].

Data extraction and quality assessment

Data was extracted by one author (RN) and checked by another (RD). Disagreements were resolved through discussion with the review team (AG, AT, LS, and RD). In instances of missing or inconsistent data, study authors were contacted. Where there was no response, data was documented as not reported, or the paper excluded. In the event of multiple publications reporting identical or heavily overlapping study populations (e.g., same registry, cancer, stage, age group, and time period), data was extracted from the earliest publication, and where there was more than one publication from the same year, extraction first prioritized the publication reporting an income-based socio-economic measure and, second, one reporting multiple socio-economic measures. If more than one multivariable analysis was conducted, information was extracted from the most comprehensive adjusted model.

Data was extracted on author(s); publication year; country; data source; number in study population; cancer diagnosis time frame, patient age(s), cancer stage, and registry coverage; socio-economic measure and unit; numbers receiving predictive biomarker test/biological and precision therapy, overall and by socio-economic group (numerator and, where available, denominator); comparator(s) (where appropriate); and measures of association for not receiving testing/treatment by socio-economic status (e.g., ORs, 95% CIs, and *p* values). All eligible studies were quality appraised using a modified version of the ISPOR checklist for retrospective database studies. Focus in particular was paid to data sources, statistical results of interest, and generalizability of conclusions drawn [17]. The tool had ten features each scored as 0, 0.5, or 1 (Additional file 1: Supplementary methods 4). Appraisal was conducted independently by two authors (RN and RD), with disagreements resolved through discussion with a third author (AT), and consensus (AT).

Synthesis of evidence

Data was synthesized using a summary of findings table. Where not reported, percentages utilizing biological and precision therapies and/or predictive biomarker test by socio-economic sub-group were calculated from data reported in the paper or supplied by authors, and unadjusted OR for low compared to high socio-economic status were computed for test/therapy receipt. Studies were heterogeneous in terms of outcome analyses (test/therapy receipt or non-receipt), socio-economic comparisons made, whether ORs (crude or adjusted) were reported, and the variables that any adjusted ORs were controlled for. Unadjusted ORs were therefore computed to enable inclusion of as many studies as possible in a consistent way. "Low" socio-economic status was defined as the lowest socio-economic sub-group in each article and "high" socio-economic status as the top sub-group.

Meta-analyses were performed using random-effects, Mantel-Haenszel methods. These assessed the likelihood of (i) test receipt and (ii) treatment receipt by low socio-economic status. Eligibility criteria for studies to be included in the meta-analysis were as follows: unadjusted low and high socio-economic utilization data for one measure of socio-economic status reported and an independent sampling frame (no data overlap with another study/paper). In the primary analyses, results relating to an income measure (or, failing that, education, or otherwise, the reported measure) were included. This reflected the dominance of USA studies within the evidence base, where there are cost implications for drug access [18]. Where multiple papers included study populations from the same or related databases that overlapped in terms of period of diagnosis/treatment, the publication reporting the largest total number of patients was entered into the meta-analysis.

For predictive biomarker tests, results were grouped by cancer site (breast, colorectal, lung, and melanoma). Those for biological and precision therapies were grouped by drug class (targeted therapy, biologic, and immunotherapy), while separate pre-specified sub-group analyses were conducted for breast cancer, lung cancer, and all other cancers (sub-grouped by cancer type: colorectal, head and neck, hepatobiliary, melanoma, mixed,

renal cell). A final post hoc sub-group analysis was performed for the Surveillance, Epidemiology, and End Results program (SEER) versus non-SEER registry studies. Testing for sub-group differences was computed where appropriate. Two post hoc sensitivity analyses (one involving substituting included studies with those excluded due to overlapping sampling frames and the other exploring USA versus non-USA healthcare settings) were conducted to determine the robustness of the results. The I^2 statistic was calculated to estimate the degree of statistical heterogeneity [19], and funnel plots were produced to assess publication bias in analyses of ten plus studies [20]. Statistical analyses were conducted using RevMan 5.3.

Results

Search results

The search identified 17,047 citations. After removal of duplicates, titles and abstracts of 10,722 records were screened for eligibility. After title and abstract screening, 551 records progressed to full-text review. Overall, 62 papers (reporting 58 independent studies) met the inclusion criteria (Fig. 1) and were included in the review. Eight studies reported utilization data for predictive biomarker tests [21–28], thirty-seven studies (41 papers) reported utilization data for biological and precision therapies [29–65], and 3 studies reported both [66–68]. Ten papers (Additional file 1: Table S1) had no denominator populations or only reported an average measure of socio-economic status (e.g., mean household income), and were excluded from inclusion in the meta-analysis and are not discussed further [69–78].

Study characteristics

The 48 included studies covered 7 cancers, 5 predictive biomarker tests, and 11 biological and precision therapy classifications, of which bevacizumab (12 studies) [41–45, 54–56, 58, 59, 64, 65] and trastuzumab (11 studies) [29–39] were most common. Most studies were in the USA ($n = 42$) [21, 22, 25–35, 40–68], and a majority analyzed SEER registry data ($n = 27$) [21, 22, 25, 29–33, 41, 42, 45, 47, 49, 50, 54–59, 61–64, 66–68] (Additional file 1: Fig. S1). Of the SEER data studies, 19 [29–32, 41, 42, 47, 49, 54–59, 61–64, 68] were SEER Medicare (i.e., included patients ≥ 65). The remaining studies were from Canada (4 studies) [23, 36–38], China (1 study) [39], and Ireland (1 study) [24]. Forty-six studies reported one or more area-based socio-economic status measure, and only two utilized individual-based measures [34, 68]. Six SES measures (poverty, income, education, employment, deprivation, and socio-economic status aggregate score) were reported. For nine studies, utilization was only available as percentages [24, 29, 32, 52, 54, 56, 61, 66, 67]. Study characteristics are summarized in Additional file 1: Table S2.

Seven papers, pertaining to four studies, reported the same data from the same registry [38, 43, 45, 79–82]. Sixteen papers (covering 8 studies) overlapped in their study populations (cancer site, stage, years of diagnosis time frames, patients' age) [29–32, 36, 37, 41–44, 49, 50, 54, 55, 67, 68]. Two studies did not report unadjusted drug and/or test utilization data [40, 58]. This left 38 studies (including 1,036,125 patients) which were included in the meta-analysis [21–29, 31, 33–35, 37–39, 42, 44–49, 51–54, 56, 57, 59–66, 68].

Quality appraisal

The 48 studies scored in the range 4–10, out of a possible 10 (mean = 6.9, median = 6.5) (Additional file 1: Table S3). Papers scored well regarding data source(s), study populations, and reporting socio-economic definition(s). Discussion of results with reference to the role of socio-economic status, statistical analysis with summary measures like OR, and explanations for confounder selection were often reported poorly.

Predictive biomarker testing

Eleven studies reported data of interest for five predictive biomarker tests [21–28, 66–68]. Ten studies were included in the meta-analysis [21–28, 66, 68]. These covered the following cancers: breast (4 studies) [21–24], colorectal (3 studies) [25–27], melanoma (1 study) [66], and non-small cell lung (2 studies) [28, 68]. The pooled OR for predictive biomarker test receipt for low socio-economic status was 0.86 (95% CI 0.71–1.05; $I^2 = 86\%$; 10 studies) (Fig. 2). This pattern was consistent across cancer sub-groups (4 breast cancer studies, 2 lung cancer studies, and 1 melanoma study) but was only significant in colorectal cancer (0.76, 95% CI 0.65–0.88; 3 studies).

Biological and precision therapies: primary analysis

Association of socio-economic status with biological and precision therapy receipt was reported in 40 studies [29–68]. Thirty of which were included in the meta-analysis [29, 31, 33–35, 37–39, 42, 44–49, 51–54, 56, 57, 59–66, 68]. The overall pooled OR for receipt of biological and precision therapy for patients from low socio-economic status was 0.83 (95% CI 0.75–0.91; $I^2 = 85\%$; 30 studies) (Fig. 3). Sub-group analysis suggested stronger associations with immunotherapy utilization (0.82, 95% CI 0.78–0.86; 7 studies) than other therapy classes (14 targeted therapy and 9 biological therapy studies), but the test for sub-group differences was not significant (Fig. 3). Sensitivity analyses which substituted included studies for excluded studies with overlapping sampling frames confirmed the robustness of results

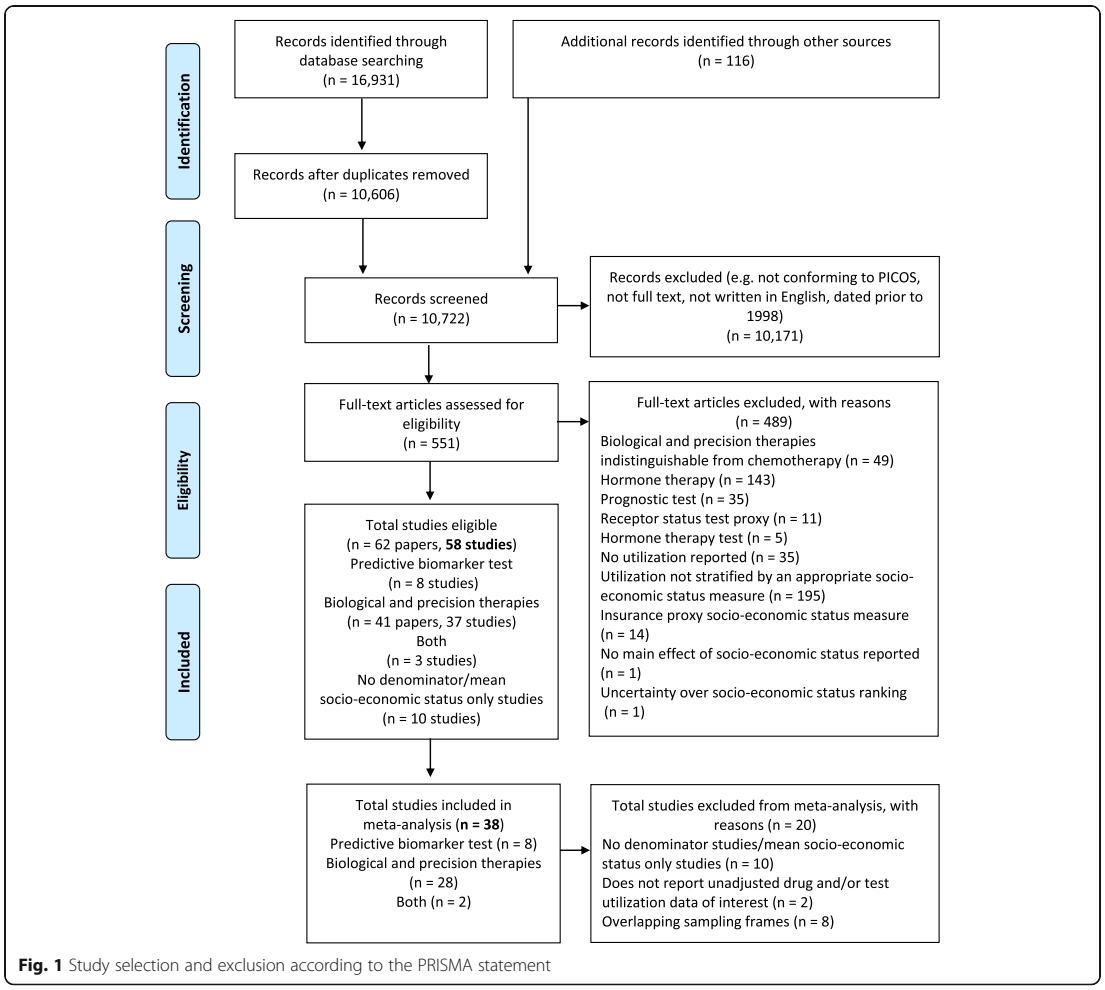


Fig. 1 Study selection and exclusion according to the PRISMA statement

(0.80, 95% CI 0.72–0.88; I^2 = 86%; 30 studies). Similar results were also observed in sensitivity analyses when only USA studies were considered (0.82, 95% CI 0.74–0.91, I^2 = 85%, 27 studies). For full sensitivity analyses results, see Additional File 1: Fig. S2.

Biological and precision therapies: sub-group analyses

For breast cancer, 11 studies reported the association of socio-economic status with the human epidermal growth factor receptor 2 (HER2) targeting monoclonal antibody trastuzumab [29–39] and one with immunotherapy [40]. Eight studies were eligible for meta-analysis [29, 31, 33–35, 37–39]. The pooled OR for receipt of trastuzumab in those with low compared to high socio-economic status was 0.93 (95% CI 0.78–1.10; I^2 = 68%) (Fig. 4).

Nine lung cancer studies evaluated socio-economic status with biological and precision therapy receipt [41–47, 67, 68]. Four of these reported bevacizumab [41–44], 2 tyrosine kinase inhibitors [67, 68], 1 both bevacizumab and tyrosine kinase inhibitors [45], 1 immunotherapy [46], and 1 biological therapies (mostly bevacizumab) [47]. Six were eligible for meta-analysis [42, 44–47, 68], and the pooled OR for receipt of biological and precision therapies in those of low compared to high socio-economic status was 0.71 (95% CI 0.51–1.00; I^2 = 95%) (Fig. 5).

Twenty studies reported data of interest for 6 other cancers: hepatobiliary (4 studies) [48–51], melanoma (3 studies) [52, 53, 66], colorectal (8 studies) [40, 54–60], renal cell carcinoma (1 study) [61], and head and neck cancer (2 studies) [62, 63]. A further two studies reported data on more than one cancer [64, 65]. Studies referenced the following 7 treatments: immunotherapy [40, 48, 52, 53, 60], bevacizumab [54–56, 58, 59, 64, 65],

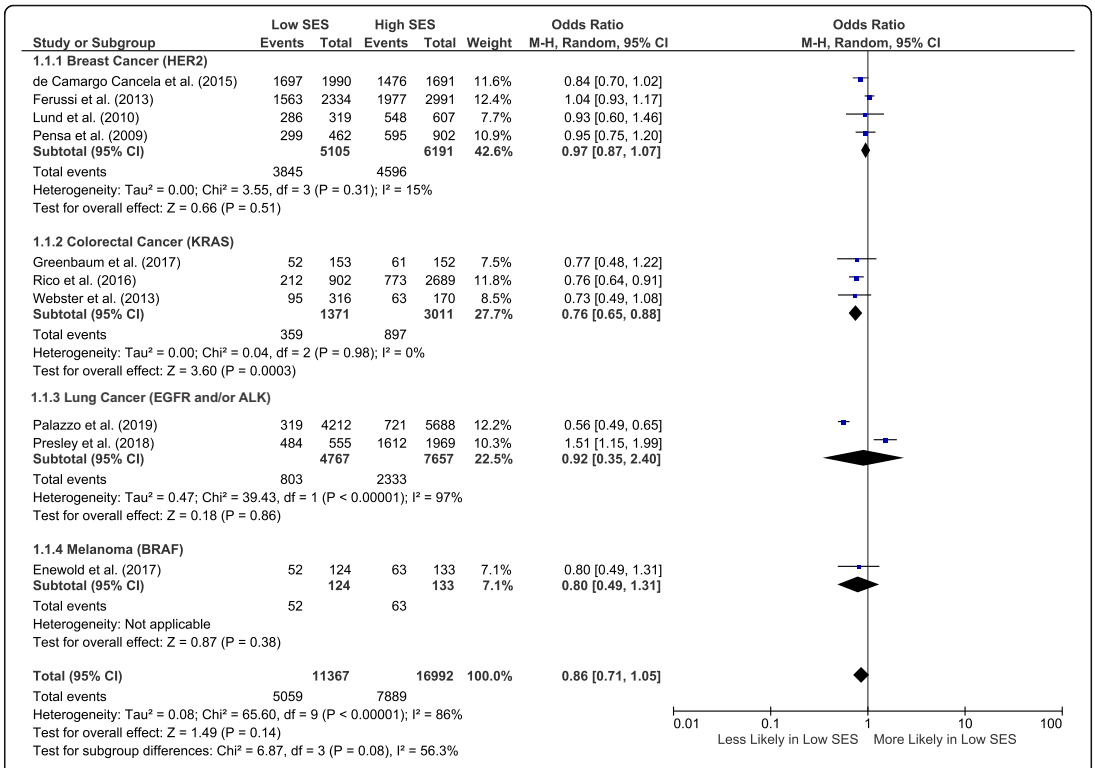


Fig. 2 Forest plot showing predictive biomarker test utilization odds (sub-grouped by cancer type) for low compared to high socio-economic status. ALK, anaplastic lymphoma kinase; BRAF, proto-oncogene B-Raf; HER2, human epidermal growth factor receptor 2; KRAS, oncogene KRAS; SES, socio-economic status

sorafenib [49–51], ipilimumab [66], targeted biologics [57], IL-2 [61], and cetuximab [62, 63]. Sixteen studies could be combined into meta-analyses [48, 49, 51–54, 56, 57, 59–66], giving a pooled OR for receipt of biological and precision therapies for low socio-economic status of 0.84 (95% CI 0.76–0.94; $I^2 = 73\%$) (Additional file 1: Fig. S3). The test for sub-group differences between breast, lung, and all other cancers was not significant (Additional file 1: Fig. S4).

Discussion

This is the first systematic review and meta-analysis to examine whether there are inequalities in novel cancer therapeutics and/or associated testing use. Overall, the findings show that there are statistically significant socio-economic inequalities in biological and precision therapy utilization; those with a low socio-economic status were 17% less likely to be treated with precision therapies. An effect of similar magnitude was observed in test receipt, but did not achieve statistical significance.

The finding that differences are present in novel cancer treatments is consistent with previous systematic reviews documenting traditional treatment inequalities [2, 3]. Similar socio-economic inequalities have also been observed across the cancer care pathway (from screening [83], to diagnosis [84], and timeliness of referral and treatment receipt [85] through to survival [86]). Combined, this suggests that low socio-economic status remains a barrier to treatment access and cancer care, despite advances in treatment.

The strength of socio-economic inequalities varied with cancer type: the effect estimate for receipt of biological and precision therapies was stronger for lung cancer (incidence of which is related to low socio-economic status) than other cancers. It is not clear why this is so, although the risk of some cancers (including lung) is associated with health behaviors (e.g., smoking) [87]. It is possible that these health behaviors, alongside other factors (which, themselves may be a consequence of the health behaviors), such as multi-morbidity, could influence a healthcare professional's decision to offer or

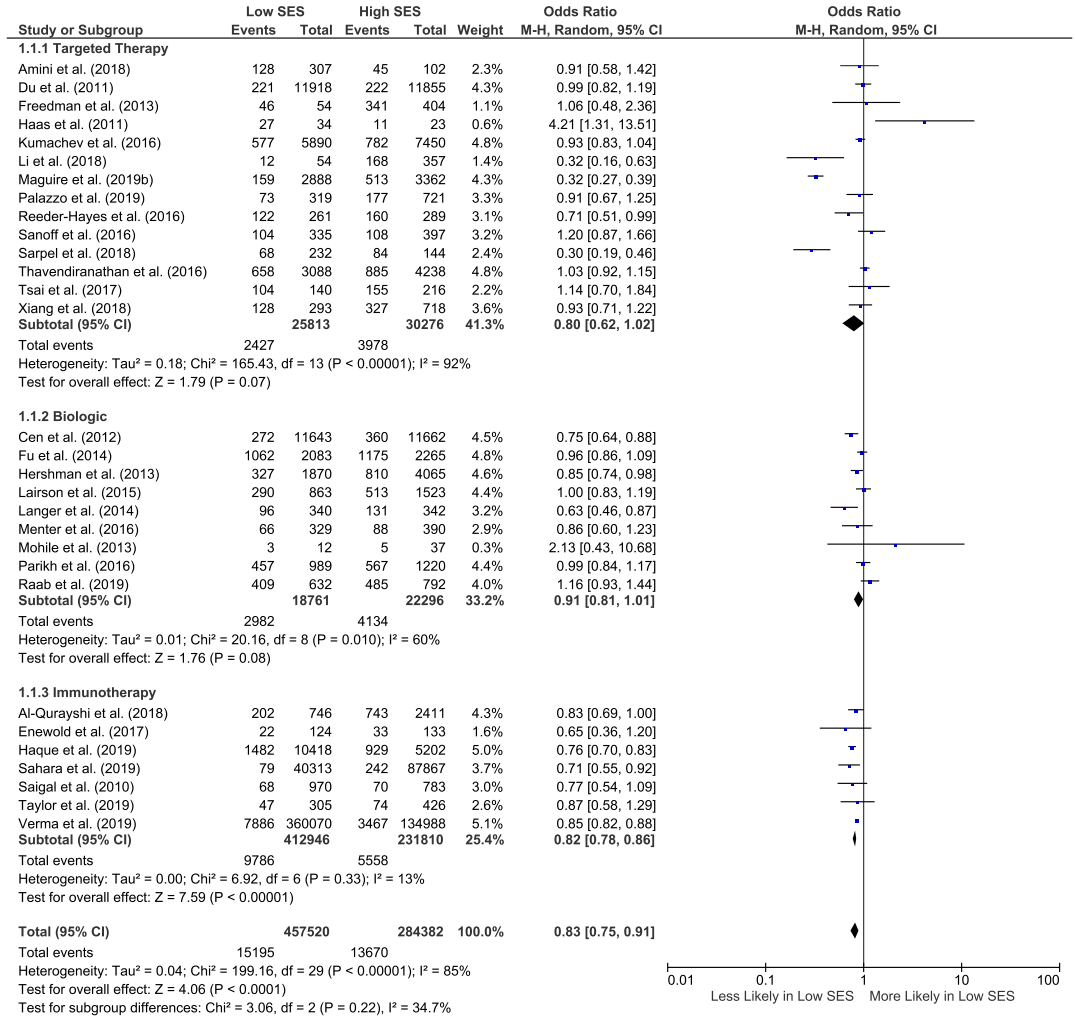


Fig. 3 Forest plot showing biological and precision therapy utilization odds for all cancers (sub-grouped by drug class) for low compared to high socio-economic status. CI, confidence interval; SES, socio-economic status

initiate, or a patient's choice to receive, cancer treatment [88]. While such individual behavioral factors warrant further investigation, they need contextualizing within the wider determinants of health (i.e., the social, economic, cultural, and clinical level factors) which are also associated with known treatment barriers [89].

The socio-economic inequalities in testing and therapy utilization in breast cancer were less pronounced, despite the majority of research focusing on this cancer. This finding, along with a previous systematic review concluding equivocal associations between socio-economic status and trastuzumab uptake [90], suggests that low socio-economic status may be less of a treatment barrier in breast cancer, at least as far as newer therapies are concerned. One possible explanation for this may be that breast cancer sub-type differentiation and the practice of hormone receptor status testing, and basing treatment on these results, are well established and routinely embedded in clinical practice (originating in the 1970s following the discovery of the estrogen receptor) [91]. Hence, our findings support the wider concept of the inverse equity hypothesis [92]: that is, that while new interventions may temporarily widen inequalities by disproportionately favoring those with resources

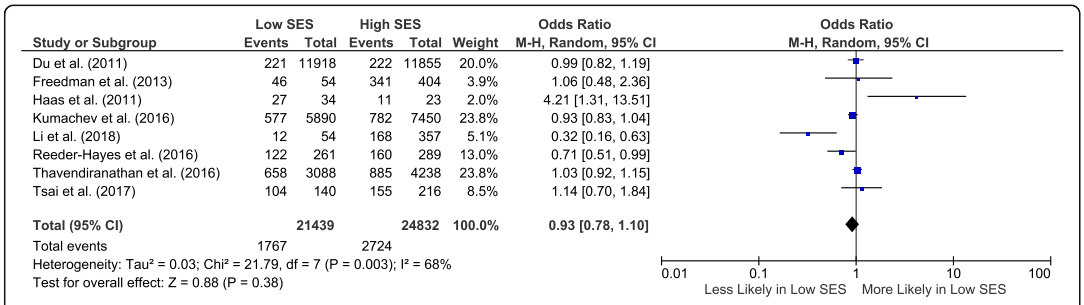


Fig. 4 Forest plot showing biological and precision therapy utilization odds in breast cancer for low compared to high socio-economic status. CI, confidence interval; SES, socio-economic status

enabling priority access, over time this narrows as treatment access "trickles down" and becomes standard clinical practice [93, 94].

In relation to predictive biomarker testing more generally, the observation that there is reduced utilization with respect to socio-economic status builds on previously documented relationships between factors associated with socio-economic status and test receipt (e.g., negative association between smoking and epidermal growth factor receptor (EGFR) and anaplastic lymphoma kinase (ALK) abnormalities) [95]. Previous work also highlights that test patterns vary temporally and spatially [96, 97], as well as with respect to patient demographics (e.g., age) [98]. This suggests that testing access is complex. Nevertheless, the observation that low socio-economic status may reduce access to testing has important implications. First, utilization barriers occur at points other than just therapy receipt, a finding echoed by Cancer Research UK who highlighted that many colorectal and non-small cell lung cancer patients potentially eligible for targeted treatments did not receive molecular testing [99]. Second, if multiple barriers to novel therapy utilization exist, then sophisticated solutions are likely required to prevent cancer inequalities widening further. In the first instance, further monitoring of inequalities is required. However, given the rapidly evolving nature of the precision oncology field, and the fact that routine datasets generally lack good data on newly licensed therapies and tests, an appreciation of how such information might be captured in future observational studies is required, especially those that are large-scale and population-based. Analysis of new data sources, rich in biological and precision therapies (e.g., UK's Systemic Anti-cancer Therapy dataset) or predictive biomarker test information (e.g., USA's Flatiron Health electronic healthcare records database), may provide the first steps here. Encouragement of data collection to enable audit of treatment access would inform development of solutions to respond to any inequalities noted (e.g., low-income assistance programs, investigating access barriers in problem areas) [9].

This is the first comprehensive meta-analysis on this important and growing area of practice, and brings together data on over 1 million patients. Despite this, the study does have several limitations. First, there are challenges comparing studies reporting different measures of socio-economic status. There was no one consistent measure used, and even when studies appeared

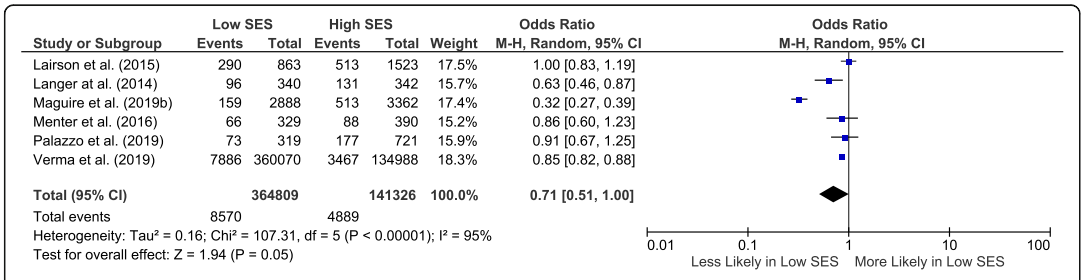


Fig. 5 Forest plot showing biological and precision therapy utilization odds in lung cancer for low compared to high socio-economic status. CI, confidence interval; SES, socio-economic status

to use the same measure (e.g., income), how the variable was categorized (e.g., what was considered "high"; number of sub-groups considered) differed. For most studies, there was considerable variation between what was classified as "high" and "low" socio-economic status, meaning that true differences were unlikely to be attenuated by a lack of variability. However, almost all studies used area-based socio-economic measures, so the ecological fallacy in inference is a risk. Secondly, determining OR from raw data disregards adjustments for confounders; this along with variations in study sampling frames may in part explain the high heterogeneity observed. It also means that the possibility cannot be entirely excluded that any associations seen in the meta-analyses could be explained by uncontrolled confounding. Third, single reviewer title and abstract screening, while considered acceptable by the Cochrane Collaboration [100], may have erroneously excluded relevant studies. Finally, any conclusions drawn here are time specific and may not fully reflect all inequalities present within the system.

The review also highlights limitations in the evidence base. For example, sub-group analyses require care in interpretation where study numbers are small. The majority of studies reported data from non-universal healthcare systems and recorded in SEER Medicare registries. As the relevance of socio-economic indicators varies across the life course, measures such as median household income may be less meaningful in retired SEER populations [101]. In such circumstances, eligibility for Medicare may be more important in addressing one of the most important barriers to care in the USA—that of having health insurance. Similarly, as employment is often tied to insurance coverage in non-universal healthcare systems like the USA, this choice of socio-economic indicator could be an additional factor related to utilization outcomes in the under 65 age group other than income alone. The generalizability of conclusions drawn to patients outside the USA and age groups younger than 65 years must be questioned. Having said this, studies from other countries documented similar patterns in inequality [24, 37, 39]. Moving forward, consideration of data from other registries (e.g., Scandinavian datasets known to be rich in socio-economic detail) would be valuable. The SEER registry also underrepresents minority populations. This limitation may be important given the links between ethnicity and genetics [102–104]. Despite these limitations, among all analyses, there was no clear observable evidence of publication bias (Additional file 1: Fig. S5).

Future research should focus on investigating the reasons for inequalities around these novel therapies. Consideration of testing as a treatment barrier requires prioritization, and work investigating clinician, patient, and family roles in decision-making around testing and treatment receipt is crucial. This is even more pertinent given the projected increases in panel sequencing testing costs and the growing number of therapeutic agents entering clinical practice. To aid further work in this area, it would be helpful if researchers critically evaluated the relationships between the different measures of socio-economic status in healthcare utilization research: for example, individual versus population measures, single versus aggregate measures, or the various single measures such as education or income. Doing so acknowledges that there is not one standardized, superior socio-economic measure to select. Rather, that as all indicators have limitations and the constraints of current dataset access may restrict the feasibility of further measurement, the magnitude of inequalities observed as well as the ability to make cross-study comparisons requires contextualizing in any future findings. From a practice perspective, policymakers and clinicians need to be aware of the potential barriers to biological and precision therapy beyond patients' tumor molecular profiles. Revising guidelines to include a focus on reducing inequalities would assist with such prioritization.

Conclusions

There are socio-economic inequalities in the utilization of both predictive biomarker tests as well as biological and precision cancer therapies. This requires further investigation to prevent differences in outcomes due to inequalities in treatment with biological and precision therapies.

Supplementary information

Supplementary information accompanies this paper at https://doi.org/10.1186/s12916-020-01753-0.

Additional file 1: Supplementary methods 1, PRISMA checklist. **Supplementary methods 2**, search strategy. **Supplementary methods 3**, inclusion/exclusion decision trees. **Supplementary methods 4**, quality appraisal tool. **Table S1**, no denominator/mean socio-economic status only study characteristics. **Fig. S1**, SEER versus non-SEER registry studies forest plot and funnel plot. **Table S2**, included studies characteristics. **Table S3**, quality appraisal results breakdown. **Fig. S2**, sensitivity analyses. **Fig. S3**, all other cancers forest plot and funnel plot. **Fig. S4**, breast, lung and all other cancers forest plot and funnel plot. **Fig. S5**, additional funnel plots.

Abbreviations

ALK: Anaplastic lymphoma kinase; BRAF: Proto-oncogene B-Raf; CI/CIs: 95% Confidence interval; EGFR: Epidermal growth factor receptor; HER2: Human epidermal growth factor receptor 2; KRAS: Oncogene KRAS; OR/ORs: Odds ratio; PICOS: Population, Intervention, Comparison, Outcome, and Setting; PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-analyses; SEER: Surveillance, Epidemiology, and End Results program; SES: Socio-economic status

Acknowledgements

With special thanks to Alex Inskip at the Walton Library, Newcastle University, for his advice in developing the search strategy.

Authors' contributions
AG, AT, and LS obtained the funding. AG, AT, LS, and RN conceived the study concept and design. AT, RD, and RN identified the literature and abstracted data. AG, AT, KJ, LS, RD, and RN interpreted the data. RN and SR undertook the statistical analysis. RN drafted the manuscript. All authors read and approved the final draft. AT has full access to all the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.

Authors' information
Not applicable

Funding
This work was funded by a Barbour Foundation PhD Studentship. The funding source had no role in the design and conduct of the study; the collection, management, analysis, and interpretation of the data; the preparation, review, or approval of the manuscript; and the decision to submit the manuscript for publication.

Availability of data and materials
All data generated or analyzed during this study are included in this published article and its supplementary information files.

Ethics approval and consent to participate
Not applicable

Consent for publication
Not applicable

Competing interests
The authors declare that they have no competing interests.

Author details
[1]School of Pharmacy, Newcastle University, King George VI Building, King's Road, Newcastle-upon-Tyne NE1 7RU, UK. [2]Population Health Sciences Institute, Newcastle University Centre for Cancer, Newcastle-upon-Tyne, UK. [3]Newcastle University Centre for Cancer, Newcastle-upon-Tyne, UK. [4]Health Economics Group and Evidence Synthesis Team, Population Health Sciences Institute, Newcastle University, Newcastle-upon-Tyne, UK. [5]Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, Stockholm, Sweden.

Received: 22 April 2020 Accepted: 19 August 2020
Published online: 23 October 2020

References

1. Macmillan. Health inequalities: time to talk. 2019. https://macmillan.org.uk/assets/health-inequalities-paper-april-2019.pdf. Accessed 15 Mar 2020.
2. Forrest LF, Adams J, Wareham H, Rubin G, White M. Socioeconomic inequalities in lung cancer treatment: systematic review and meta-analysis. PLoS Med. 2013;10(2):e1001376-e.
3. Aarts MJ, Lemmens VEPP, Louwman MWJ, Kunst AE, Coebergh JWW. Socioeconomic status and changing inequalities in colorectal cancer? A review of the associations with risk, treatment and outcome. Eur J Cancer. 2010;46(15):2681–95.
4. Halpern MT. Cancer disparities research: it is time to come of age. Cancer. 2015;121(8):1158–9.
5. Dranitsaris G, Zhu X, Adunlin G, Vincent MD. Cost effectiveness vs affordability in the age of immuno-oncology cancer drugs. Expert Rev Pharmacoecon Outcomes Res. 2018;18(4):351–7.
6. NCCN. Clinical practice guidelines in oncology: non-small cell lung cancer. 2020. http://www.nccn.org/professionals/physician_gls/pdf/nscl.pdf. Accessed 23 Mar 2020.
7. IQVIA Institute. Global oncology trends 2019 therapeutics, clinical development and health system implications institute report. 2019. https://www.iqvia.com/-/media/iqvia/pdfs/institute-reports/global-oncology-trends-2019.pdf?_=1587463035271. Accessed 14 Mar 2020.
8. McClellan KA, Avard D, Simard J, Knoppers BM. Personalized medicine and access to health care: potential for inequitable access? Eur J Hum Genet. 2013;21(2):143–7.
9. Ward MM. Personalized therapeutics: a potential threat to health equity. J Gen Intern Med. 2012;27(7):868–70.
10. James JE. Personalised medicine, disease prevention, and the inverse care law: more harm than benefit? Eur J Epidemiol. 2014;29(6):383–90.
11. Huey RW, Hawk E, Offodile AC. Mind the gap: precision oncology and its potential to widen disparities. J Onco Practice. 2019;15(6):301–4.
12. Schwaederle M, Zhao M, Lee JJ, Eggermont AM, Schilsky RL, Mendelsohn J, et al. Impact of precision medicine in diverse cancers: a meta-analysis of phase II clinical trials. J Clin Oncol. 2015;33(32):3817–25.
13. Mok TS, Wu Y-L, Thongprasert S, Yang C-H, Chu D-T, Saijo N, et al. Gefitinib or carboplatin–paclitaxel in pulmonary adenocarcinoma. N Engl J Med. 2009;361(10):947–57.
14. Moher D, Liberati A, Tetzlaff J, Altman DG, Group P. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. PLoS Med. 2009;6(7):e1000097-e.
15. Ross JS. Targeted therapies for cancer. In: Gingsburg GS, Willard HF, editors. Essentials of genomic and personalized medicine. 1st ed. Amsterdam; Boston; London: Elsevier; 2010. p. 532–547.
16. Higgins J, Deek JJ. Chapter 7: selecting studies and collecting data. In: Higgins JPT, Green S, editors. Cochrane handbook for systematic reviews of interventions. Version 510 [updated March 2011]. Chichester: Wiley; 2011.
17. Motheral B, Brooks J, Clark MA, Crown WH, Davey P, Hutchins D, et al. A checklist for retrospective database studies—report of the ISPOR task force on retrospective databases. Value Health. 2003;6(2):90–7.
18. Braveman PA, Cubbin C, Egerter S, Chideya S, Marchi KS, Metzler M, et al. Socioeconomic status in health research one size does not fit all. JAMA. 2005;294(22):2879–88.
19. Higgins JPT, Thompson SG. Quantifying heterogeneity in a meta-analysis. Stat Med. 2002;21(11):1539–58.
20. Egger M, Smith GD, Schneider M, Minder C. Bias in meta-analysis detected by a simple, graphical test. BMJ. 1997;315(7109):629–34.
21. Pensa M, Swede H, Brockmeyer JA, Gregorio DI. Patterns of HER2 testing in the management of primary breast cancer. Cancer Epidemiol. 2009;33(2):113–7.
22. Lund MJ, Butler EN, Hair BY, Ward KC, Andrews JH, Oprea-Ilies G, et al. Age/race differences in HER2 testing and in incidence rates for breast cancer triple subtypes. Cancer. 2010;116(11):2549–59.
23. Ferrusi IL, Earle CC, Trudeau M, Leighl NB, Pullenayegum E, Khong H, et al. Closing the personalized medicine information gap: HER2 test documentation practice. Am J Manag Care. 2013;19(1):838–44.
24. de Camargo CM, Comber H, Sharp L. Which women with breast cancer do, and do not, undergo receptor status testing? A population-based study. Cancer Epidemiol. 2015;39(5):778–82.
25. Greenbaum A, Wiggins C, Meisner AL, Rojo M, Kinney AY, Rajput A. KRAS biomarker testing disparities in colorectal cancer patients in New Mexico. Heliyon. 2017;(3, 11):e00448-e.
26. Rico A, Pollack LA, Thompson TD, Hsieh M-C, Wu X-C, Karlitz JJ, et al. KRAS testing and first-line treatment among patients diagnosed with metastatic colorectal cancer using population data from ten national program of cancer registries in the United States. J Cancer Res Ther (Manch). 2016;5(2):7–13.
27. Webster J, Kauffman TL, Feigelson HS, Pawloski PA, Onitilo AA, Potosky AL, et al. KRAS testing and epidermal growth factor receptor inhibitor treatment for colorectal cancer in community settings. Cancer Epidemiol Biomark Prev. 2013;22(1):91–101.
28. Presley CJ, Tang D, Soulos PR, Chiang AC, Longtine JA, Adelson KB, et al. Association of broad-based genomic sequencing with survival among patients with advanced non–small cell lung cancer in the community oncology setting. JAMA. 2018;320(5):469–77.
29. Du XL, Xia R, Burau K, Liu C-C. Cardiac risk associated with the receipt of anthracycline and trastuzumab in a large nationwide cohort of older women with breast cancer, 1998–2005. Med Oncol. 2011;28(1):80–90.
30. Vaz-Luis I, Lin NU, Keating NL, Barry WT, Lii H, Winer EP, et al. Racial differences in outcomes for patients with metastatic breast cancer by disease subtype. Breast Cancer Res Treat. 2015;151(3):697–707.
31. Reeder-Hayes K, Hinton SP, Meng K, Carey LA, Dusetzina SB. Disparities in use of human epidermal growth hormone receptor 2–targeted therapy for early-stage breast cancer. J Clin Oncol. 2016;34(17):2003–9.
32. Vaz-Luis I, Lin NU, Keating NL, Barry WT, Lii J, Burstein HJ, et al. Treatment of early-stage human epidermal growth factor 2-positive cancers among medicare enrollees: age and race strongly associated with non-use of trastuzumab. Breast Cancer Res Treat. 2016;159(1):151–62.

Appendix 2.1 Continued

33. Tsai H-T, Isaacs C, Lynce FC, O'Neill SC, Liu C, Schwartz MD, et al. Initiation of trastuzumab by women younger than 64 years for adjuvant treatment of stage I–III breast cancer. J Natl Compr Cancer Netw. 2017;15(5):601–7.
34. Freedman RA, Hughes ME, Ottesen RA, Weeks JC, He Y, Wong Y-N, et al. Use of adjuvant trastuzumab in women with human epidermal growth factor receptor 2 (HER2)-positive breast cancer by race/ethnicity and education within the national comprehensive cancer network. Cancer. 2013;119(4):839–46.
35. Haas JS, Phillips KA, Liang S-Y, Hassett MJ, Keohane C, Elkin EB, et al. Genomic testing and therapies for breast cancer in clinical practice. J Oncol Pract. 2011;7(3 Suppl):e1s–7s.
36. Goldhar HA, Yan AT, Ko DT, Earle CC, Tomlinson GA, Trudeau ME, et al. The temporal risk of heart failure associated with adjuvant trastuzumab in breast cancer patients: a population study. J Natl Cancer Inst. 2015;108(1):djv301.
37. Kumachev A, Trudeau ME, Chan KKW. Associations among socioeconomic status, patterns of care, and outcomes in breast cancer patients in a universal health care system: Ontario's experience. Cancer. 2016;122(6):893–8.
38. Thavendiranathan P, Abdel-Qadir H, Fischer HD, Camacho X, Amir E, Austin PC, et al. Breast cancer therapy–related cardiac dysfunction in adult women treated in routine clinical practice: a population-based cohort study. J Clin Oncol. 2016;34(19):2239–46.
39. Li J, Shao Z, Xu B, Jiang Z, Cui S, Zhang J, et al. Use of trastuzumab as an adjuvant/neoadjuvant therapy in patients with HER2-positive breast cancer in China: the Nvwa study. Medicine (Baltimore). 2018;97(21):e10350-e.
40. Shih Y-CT, Elting LS, Halpern MT. Factors associated with immunotherapy use among newly diagnosed cancer patients. Med Care. 2009;47(9):948–58.
41. Zhu J, Sharma DB, Gray SW, Chen AB, Weeks JC, Schrag D. Carboplatin and paclitaxel with vs without bevacizumab in older patients with advanced non–small cell lung cancer. JAMA. 2012;307(15):1593–601.
42. Langer C, Ravelo A, Hazard SJ, Guerin A, Ionescu-Ittu R, Latremouille-Viau D, et al. Comparison of survival and hospitalization rates between Medicare patients with advanced NSCLC treated with bevacizumab–carboplatin–paclitaxel and carboplatin–paclitaxel: a retrospective cohort study. Lung Cancer. 2014;86(3):350–7.
43. Ritzwoller DP, Carroll NM, Delate T, Hornbrook MC, Kushi L, Bowles EJA, et al. Comparative effectiveness of adjunctive bevacizumab for advanced lung cancer: the cancer research network experience. J Thorac Oncol. 2014;9(5):692–701.
44. Menter AR, Carroll NM, Sakoda LC, Delate T, Hornbrook MC, Jain RK, et al. Effect of angiotensin system inhibitors on survival in patients receiving chemotherapy for advanced non–small-cell lung cancer. Clin Lung Cancer. 2017;18(2):189–97 e3.
45. Maguire FB, Morris CR, Parikh-Patel A, Cress RD, Keegan THM, Li C-S, et al. First-line systemic treatments for stage IV non-small cell lung cancer in California: patterns of care and outcomes in a real-world setting. JNCI Cancer Spectrum. 2019b;3(3):pkz020.
46. Verma V, Haque W, Cushman TR, Lin C, II CBS, Chang JY, et al. Racial and insurance-related disparities in delivery of immunotherapy-type compounds in the United States. J Immunother. 2019;42(2):55–64.
47. Lairson DR, Parikh RC, Cormier JN, Chan W, Du XL. Cost-utility analysis of treatments for advanced non-small cell lung cancer. Am J Pharm Benefits. 2015;7(6):271–9.
48. Sahara K, Farooq SA, Tsilimigras DI, Merath K, Paredes AZ, Wu L, et al. Immunotherapy utilization for hepatobiliary cancer in the United States: disparities among patients with different socioeconomic status. Hepatobiliary Surg Nutr. 2019;9(1):13–24.
49. Sanoff HK, Chang Y, Lund JL, O'Neil BH, Dusetzina SB. Sorafenib effectiveness in advanced hepatocellular carcinoma. Oncologist. 2016;21(9):1113–20.
50. Parsons HM, Chu Q, Karlitz JJ, Stevens JL, Harlan LC. Adoption of sorafenib for the treatment of advanced-stage hepatocellular carcinoma in oncology practices in the United States. Liver Cancer. 2017;6(3):216–26.
51. Sarpel U, Heskel M, Spivack JH, Feferman Y, Ang C, Gany F. Disparities in access to sorafenib in communities with low socioeconomic status. J Health Care Poor Underserved. 2018;29(3):1123–34.
52. Al-Qurayshi Z, Crowther JE, Hamner JB, Ducoin C, Killackey MT, Kandil E. Disparities of immunotherapy utilization in patients with stage III cutaneous melanoma: a national perspective. Anticancer Res. 2018;38(5):2897–901.
53. Haque W, Verma V, Butler EB, Teh BS. Racial and socioeconomic disparities in the delivery of immunotherapy for metastatic melanoma in the United States. J Immunother. 2019;42(6):228–35.
54. Fu AZ, Tsai H-T, Marshall JL, Freedman AN, Potosky AL. Utilization of bevacizumab in US elderly patients with colorectal cancer receiving chemotherapy. J Oncol Pharm Pract. 2014;20(5):332–40.
55. Meyerhardt JA, Li L, Sanoff HK, Carpenter W, Schrag D. Effectiveness of bevacizumab with first-line combination chemotherapy for Medicare patients with stage IV colorectal cancer. J Clin Oncol. 2012;30(6):608–15.
56. Cen P, Liu C, Du XL. Comparison of toxicity profiles of fluorouracil versus oxaliplatin regimens in a large population-based cohort of elderly patients with colorectal cancer. Ann Oncol. 2012;23(6):1503–11.
57. Parikh RC, Du XL, Morgan RO, Lairson DR. Patterns of treatment sequences in chemotherapy and targeted biologics for metastatic colorectal cancer: findings from a large community-based cohort of elderly patients. Drugs Real World Outcomes. 2016;3(1):69–82.
58. Neugut AI, Becker DJ, Insel BJ, Hershman DL. Uptake of oxaliplatin and bevacizumab for treatment of node-positive and metastatic colon cancer. J Oncol Pract. 2012;8(3):156–63.
59. Raab GT, Lin A, Hillyer GC, Keller D, O'Neil DS, Accordino MK, et al. Use of bevacizumab for elderly patients with stage IV colon cancer: analysis of SEER-Medicare data. Clin Colorectal Cancer. 2019;18(3):e294–e9.
60. Taylor JP, Stem M, Yu D, Chen SY, Fang SH, Gearhart SL, et al. Treatment strategies and survival trends for anorectal melanoma: is it time for a change? World J Surgery. 2019;43(7):1809–19.
61. Saigal CS, Deibert CM, Lai J, Schonlau M. Disparities in the treatment of patients with IL-2 for metastatic renal cell carcinoma. Urol Oncol. 2010;28(3):308–13.
62. Amini A, Eguchi M, Jones BL, Stokes WA, Gupta A, McDermott JD, et al. Comparing outcomes of concurrent chemotherapy regimens in patients 65 years old or older with locally advanced oropharyngeal carcinoma. Cancer. 2018;124(22):4322–31.
63. Xiang M, Holsinger FC, Colevas AD, Chen MM, Le Q-T, Beadle BM. Survival of patients with head and neck cancer treated with definitive radiotherapy and concurrent cisplatin or concurrent cetuximab: a Surveillance, Epidemiology, and End Results-Medicare analysis. Cancer. 2018;124(23):4486–94.
64. Hershman DL, Wright JD, Lim E, Buono DL, Tsai WY, Neugut AI. Contraindicated use of bevacizumab and toxicity in elderly patients with cancer. J Clin Oncol. 2013;31(28):3592–9.
65. Mohile SG, Hardt M, Tew W, Owusu C, Klepin H, Gross C, et al. Toxicity of bevacizumab in combination with chemotherapy in older patients. Oncologist. 2013;18(4):408–14.
66. Enewold L, Sharon E, Harlan LC. Metastatic melanoma: treatment and survival in the US after the introduction of ipilimumab and vemurafenib. Oncology Res Treat. 2017;40(4):174–83.
67. Enewold L, Thomas A. Real-world patterns of EGFR testing and treatment with erlotinib for non-small cell lung cancer in the United States. PLoS One. 2016;11(6):e0156728.
68. Palazzo LL, Sheehan DF, Tramontano AC, Kong CY. Disparities and trends in genetic testing and erlotinib treatment among metastatic non–small cell lung cancer patients. Cancer Epidemiology Biomarkers Prev. 2019;28(5):926–34.
69. Freedman RA, Vaz-Luis I, Barry WT, Lii H, Lin NU, Winer EP, et al. Patterns of chemotherapy, toxicity, and short-term outcomes for older women receiving adjuvant trastuzumab-based therapy. Breast Cancer Res Treat. 2014;145(2):491–501.
70. Vaz-Luis I, Keating NL, Lin NU, Lii H, Winer EP, Freedman RA. Duration and toxicity of adjuvant trastuzumab in older patients with early-stage breast cancer: a population-based study. J Clin Oncol. 2014;32(9):927–34.
71. Reeder-Hayes KE, Meyer AM, Hinton SP, Meng K, Carey LA, Dusetzina SB. Comparative toxicity and effectiveness of trastuzumab-based chemotherapy regimens in older women with early-stage breast cancer. J Clinic Oncol. 2017;35(29):3298–305.
72. Chavez-MacGregor M, Niu J, Zhang N, Elting LS, Smith BD, Banchs J, et al. Cardiac monitoring during adjuvant trastuzumab-based chemotherapy among older patients with breast cancer. J Clin Oncol. 2015;33(19):2176–83.
73. Lu CY, Srasuebkul P, Drew AK, Chen K, Ward RL, Pearson S-A. Trastuzumab therapy in Australia: which patients with HER2+ metastatic breast cancer are assessed for cardiac function? Breast. 2013;22(4):482–7.
74. Krimphove MJ, Tully KH, Friedlander DF, Marchese M, Ravi P, Lipsitz SR, et al. Adoption of immunotherapy in the community for patients diagnosed with metastatic melanoma. J Immunother Cancer. 2019;7(1):289.
75. Zheng Z, Hanna N, Onukwugha E, Reese ES, Seal B, Mullins CD. Does the type of first-line regimens influence the receipt of second-line chemotherapy treatment? An analysis of 3,211 metastatic colon cancer patients. Cancer Med. 2014;3(1):124–33.

76. Kwan SW, Harris WP, Gold LS, Hebert PL. Comparative effectiveness of transarterial embolization and sorafenib for hepatocellular carcinoma: a population-based study. Am J Roentgenol. 2018;210(6):1359–65.
77. Li P, Jahnke J, Pettit AR, Wong Y-N, Doshi JA. Comparative survival associated with use of targeted vs nontargeted therapy in medicare patients with metastatic renal cell carcinoma. JAMA Netw Open. 2019;2(6): e195806-e.
78. Zandberg DP, Cullen K, Bentzen SM, Goloubeva OG. Definitive radiation with concurrent cetuximab vs. radiation with or without concurrent cytotoxic chemotherapy in older patients with squamous cell carcinoma of the head and neck: analysis of the SEER-medicare linked database. Oral Oncol. 2018;86:132–40.
79. Thavendiranathan P, Abdel-Qadir H, Fischer HD, Liu Y, Camacho X, Amir E, et al. Risk-imaging mismatch in cardiac imaging practices for women receiving systemic therapy for early-stage breast cancer: a population-based cohort study. J Clin Oncol. 2018;36(30):2980–7.
80. Delate T, Won K, Nm C, Kushi L, Hornbrook M, Eja B, et al. Factors associated with first-line bevacizumab use in advanced non-squamous non-small cell lung cancer. J Cancer Res Ther (Manch). 2014;2(1):1–8.
81. Carroll NM, Delate T, Menter A, Hornbrook MC, Kushi L, Bowles EJA, et al. Use of bevacizumab in community settings: toxicity profile and risk of hospitalization in patients with advanced non–small-cell lung cancer. J Oncol Pract. 2015;11(5):356–62.
82. Maguire FB, Morris CR, Parikh-Patel A, Cress RD, Keegan THM, Li C-S, et al. Disparities in systemic treatment use in advanced-stage non–small cell lung cancer by source of health insurance. Cancer Epidemiol Biomarkers Prev. 2019a;28(6):1059–66.
83. Smith D, Thomson K, Bambra C, Todd A. The breast cancer paradox: a systematic review of the association between area-level deprivation and breast cancer screening uptake in Europe. Cancer Epidemiol. 2019;60:77–85.
84. Lyratzopoulos G, Abel GA, Brown CH, Rous BA, Vernon SA, Roland M, et al. Socio-demographic inequalities in stage of cancer diagnosis: evidence from patients with female breast, lung, colon, rectal, prostate, renal, bladder, melanoma, ovarian and endometrial cancer. Ann Oncol. 2013;24(3):843–50.
85. Hayes L AJ, McCallum I, Forrest L, Hidajat M, White M, Sharp L. Age-related and socio-economic inequalities in timeliness of referral and start of treatment in colorectal cancer. A population-based analysis J Epi Comm Health 2020; in press.
86. Woods LM, Rachet B, Coleman MP. Origins of socio-economic inequalities in cancer survival: a review. Ann Oncol. 2006;17(1):5–19.
87. Network NCI. Evidence to March 2010 on cancer inequalities in England. 2010. www.ncin.org.uk/inequalities. Accessed 17 Mar 2020.
88. Sarfati D, Koczwara B, Jackson C. The impact of comorbidity on cancer and its treatment. CA Cancer J Clin. 2016;66(4):337–50.
89. Dahlgren G, Whitehead M. Tackling inequalities in health: what can we learn from what has been tried? Working paper prepared for the King's Fund international seminar on tackling inequalities in health, September 1993, Ditchley Park, Oxfordshire. London, King's Fund. In: Dahlgren G, Whitehead M, editors. European strategies for tackling social inequities in health: Levelling up part 2 Copenhagen: WHO Regional Office for Europe 2007. p. 20–32.
90. Martin AP, Downing J, Cochrane M, Collins B, Francis B, Haycox A, et al. Trastuzumab uptake in HER2-positive breast cancer patients: a systematic review and meta-analysis of observational studies. Crit Rev Oncol/Hematol. 2018;130:92–107.
91. Osbourne CK. Steroid hormone receptors in breast cancer management. Breast Cancer Res Treat. 1998;51:227–38.
92. Victora CG, Vaughan JP, Barros FC, Silva AC, Tomasi E. Explaining trends in inequities: evidence from Brazilian child health studies. Lancet. 2000; 356(9235):1093–8.
93. Rodgers E. Diffusion of innovations. 5th ed. New York: Somon & Schuster International; 2003.
94. White M, Adams J, Heywood P. How and why do interventions that increase health overall widen inequalities within populations? In: Babones SJ, editor. Social inequality and public health. Bristol: Policy Press; 2012.
95. Tseng C-H, Chiang C-J, Tseng J-S, Yang T-Y, Hsu K-H, Chen K-C, et al. EGFR mutation, smoking, and gender in advanced lung adenocarcinoma. Oncotarget. 2017;8(58):98384–93.
96. Lung Cancer Coalition. Molecules matter turning the science of molecular diagnostics in lung cancer into a practical service for all patients. 2019. https://www.uklcc.org.uk/wp-content/uploads/2019/01/Molecules-Matter.pdf. Accessed 15 Mar 2020.
97. Lee DH, Tsao M-S, Kambartel K-O, Isobe H, Huang M-S, Barrios CH, et al. Molecular testing and treatment patterns for patients with advanced non-small cell lung cancer: PIvOTAL observational study. PloS One. 2018;13(8): e0202865-e.
98. Illei PB, Wong W, Wu N, Chu L, Gupta R, Schulze K, et al. ALK testing trends and patterns among community practices in the United States. JCO Precis Oncol. 2018;2:1–11.
99. Cancer Research UK. Cancer Research UK policy statement: patient access to molecular diagnostics and targeted medicines in England. 2018. https://www.cancerresearchuk.org/sites/default/files/access_to_molecular_diagnostic_tests_and_targeted_medicines_in_england_0.pdf. Accessed 15 Mar 2020.
100. Deeks JJ, Higgins JPT, Altman DG. Chapter 10: analysing data and undertaking meta-analyses, Section 4.6.4 In: Higgins JPT, Green S, editors. Cochrane handbook for systematic reviews of interventions. Version 510 [updated March 2011]. Chichester: Wiley; 2011.
101. Galobardes B, Shaw M, Lawlor DA, Lynch JW, Davey SG. Indicators of socioeconomic position (part 1). Journal Epidemiol Community Health. 2006;60(1):7–12.
102. Bach PB, Guadagnoli E, Schrag D, Schussler N, Warren JL. Patient demographic and socioeconomic characteristics in the SEER-Medicare database: applications and limitations. Med Care. 2002;40(8):IV19–25.
103. Griggs JJ. Role of nonclinical factors in the receipt of high-quality systemic adjuvant breast cancer treatment. J Clinical Oncol. 2012;30(2):121–4.
104. Goss E, Lopez AM, Brown CL, Wollins DS, Brawley OW, Raghavan D. American Society of Clinical Oncology policy statement: disparities in cancer care. J Clin Oncol. 2009;27(17):2881–5.

Publisher's Note

Appendix 2.2 PRISMA Checklist for Chapter 2

Section/topic	#	Checklist item	Reported on page #
TITLE			
Title	1	Identify the report as a systematic review, meta-analysis, or both.	38
ABSTRACT			
Structured summary	2	Provide a structured summary including, as applicable: background; objectives; data sources; study eligibility criteria, participants, and interventions; study appraisal and synthesis methods; results; limitations; conclusions and implications of key findings; systematic review registration number.	NA was needed for publication
INTRODUCTION			
Rationale	3	Describe the rationale for the review in the context of what is already known.	39
Objectives	4	Provide an explicit statement of questions being addressed with reference to participants, interventions, comparisons, outcomes, and study design (PICOS).	41 & 42
METHODS			
Protocol and registration	5	Indicate if a review protocol exists, if and where it can be accessed (e.g., Web address), and, if available, provide registration information including registration number.	NA
Eligibility criteria	6	Specify study characteristics (e.g., PICOS, length of follow-up) and report characteristics (e.g., years considered, language, publication status) used as criteria for eligibility, giving rationale.	41 & 42
Information sources	7	Describe all information sources (e.g., databases with dates of coverage, contact with study authors to identify additional studies) in the search and date last searched.	40
Search	8	Present full electronic search strategy for at least one database, including any limits used, such that it could be repeated.	40 Appendix 2.3
Study selection	9	State the process for selecting studies (i.e., screening, eligibility, included in systematic review, and, if applicable, included in the meta-analysis).	40, 41 & 42 Appendices 2.5 & 2.7
Data collection process	10	Describe method of data extraction from reports (e.g., piloted forms, independently, in duplicate) and any processes for obtaining and confirming data from investigators.	42 & 43

Appendix 2.1 Continued

Section/topic	#	Checklist item	Reported on page #
Data items	11	List and define all variables for which data were sought (e.g., PICOS, funding sources) and any assumptions and simplifications made.	41 & 42
Risk of bias in individual studies	12	Describe methods used for assessing risk of bias of individual studies (including specification of whether this was done at the study or outcome level), and how this information is to be used in any data synthesis.	45 & 46
Summary measures	13	State the principal summary measures (e.g., risk ratio, difference in means).	44
Synthesis of results	14	Describe the methods of handling data and combining results of studies, if done, including measures of consistency (e.g., I2) for each meta-analysis.	43, 44, 45 & 46
Risk of bias across studies	15	Specify any assessment of risk of bias that may affect the cumulative evidence (e.g., publication bias, selective reporting within studies).	45 & 46 Appendices 2.9b, 2.12b, 2.13b & 2.14
Additional analyses	16	Describe methods of additional analyses (e.g., sensitivity or subgroup analyses, meta-regression), if done, indicating which were pre-specified.	45
RESULTS			
Study selection	17	Give numbers of studies screened, assessed for eligibility, and included in the review, with reasons for exclusions at each stage, ideally with a flow diagram.	46 Figure 2.1
Study characteristics	18	For each study, present characteristics for which data were extracted (e.g., study size, PICOS, follow-up period) and provide the citations.	48, 76, 77, 78, 80 & 81 Table 2.1 Appendix 2.8
Risk of bias within studies	19	Present data on risk of bias of each study and, if available, any outcome level assessment (see item 12).	73 Table 2.2 Appendix 2.6
Results of individual studies	20	For all outcomes considered (benefits or harms), present, for each study: (a) simple summary data for each intervention group (b) effect estimates and confidence intervals, ideally with a forest plot.	76, 77, 78, 79 & 80 Figure 2.2, 2.3, 2.4 & 2.5

Appendix 2.2 Continued

Section/topic	#	Checklist item	Reported on page #
Synthesis of results	21	Present results of each meta-analysis done, including confidence intervals and measures of consistency.	76, 77, 78, 79 & 80 Figure 2.2, 2.3, 2.4 & 2.5 Table 2.2 Appendices 2.9, 2.10, 2.11, 2.12 & 2.13
Risk of bias across studies	22	Present results of any assessment of risk of bias across studies (see Item 15).	73 Table 2.2 Appendix 2.6
Additional analysis	23	Give results of additional analyses, if done (e.g., sensitivity or subgroup analyses, meta-regression [see Item 16]).	78 Appendices 2.10, & 2.11
DISCUSSION			
Summary of evidence	24	Summarize the main findings including the strength of evidence for each main outcome; consider their relevance to key groups (e.g., healthcare providers, users, and policy makers).	82, 83, 84, 85, 86, 87, 88 &189
Limitations	25	Discuss limitations at study and outcome level (e.g., risk of bias), and at review-level (e.g., incomplete retrieval of identified research, reporting bias).	84, 85, 86 & 87
Conclusions	26	Provide a general interpretation of the results in the context of other evidence, and implications for future research.	88, 209 & 210
FUNDING			
Funding	27	Describe sources of funding for the systematic review and other support (e.g., supply of data); role of funders for the systematic review.	NA was needed for publication

From: Moher D, Liberati A, Tetzlaff J, Altman DG, The PRISMA Group (2009) (242). Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. PLoS Med 6(7): e1000097. doi:10.1371/journal.pmed1000097

Appendix 2.3 Database Search Strategy for Chapter 2

A. Medline search strategy 04/02/2019 (kept updated through to 31/12/2019)

	Search Term	Number Retrieved
1.	alectinib.mp	218
2.	ceritinib.mp.	250
3.	Crizotinib/	948
4.	crizotinib.mp.	1,426
5.	brigatinib.mp.	35
6.	Erlotinib Hydrochloride/	3,524
7.	erlotinib.mp.	5,180
8.	Gefitinib/	4,090
9.	gefitinib.mp.	5,613
10.	Afatinib/	482
11.	afatinib.mp.	736
12.	osimertinib.mp.	231
13.	Imatinib Mesylate/	9,511
14.	imatinib.mp.	12,608
15.	Lapatinib/	1,407
16.	lapatinib.mp.	2,064
17.	neratinib.mp.	137
18.	masitinib.mp.	101
19.	lenvatinib.mp.	227
20.	cabozantinib.mp.	412
21.	Sunitinib/	3,150
22.	sunitinib.mp.	4,567
23.	Axitinib/	406
24.	axitinib.mp.	637
25.	pazopanib.mp.	1,095
26.	tivozanib.mp.	71
27.	vandetanib.mp.	516
28.	nintedanib.mp.	428
29.	regorafenib.mp.	516
30.	Sorafenib/	4,170
31.	sorafenib.mp.	6,010
32.	dabrafenib.mp.	614
33.	Vemurafenib/	1,085
34.	vemurafenib.mp.	1,553
35.	encorafenib.mp.	11
36.	cobimetinib.mp.	91
37.	trametinib.mp.	581
38.	binimetinib.mp.	35
39.	Everolimus/	4,079
40.	everolimus.mp.	5,388
41.	temsirolimus.mp.	1,239
42.	abemaciclib.mp.	75
43.	palbociclib.mp.	385
44.	ribociclib.mp.	101
45.	niraparib.mp.	63
46.	olaparib.mp.	689
47.	rucaparib.mp.	118
48.	vismodegib.mp.	356
49.	sonidegib.mp.	62
50.	Cetuximab/	3,976
51.	cetuximab.mp.	5,544
52.	Panitumumab/	851
53.	panitumumab.mp.	1,327
54.	pertuzumab.mp.	647
55.	Trastuzumab/	6,015
56.	trastuzumab.mp.	8,416
57.	trastuzumab emtansine.mp.	386

Appendix 2.3 Continued

58.	exp Tamoxifen/	15,087
59.	tamoxifen.mp.	18,378
60.	Fulvestrant/	2,080
61.	fulvestrant.mp.	2,510
62.	toremifene.mp.	520
63.	Megestrol Acetate/	512
64.	megestrol acetate.mp.	771
65.	Medroxyprogesterone Acetate/	2,936
66.	medroxyprogesterone acetate.mp.	3,923
67.	Anastrozole/	1,301
68.	anastrozole.mp.	1,861
69.	exemestane.mp.	1,151
70.	Letrozole/	1,787
71.	letrozole.mp.	2,471
72.	Goserelin/	1,020
73.	goserelin.mp.	1,182
74.	Buserelin/	644
75.	buserelin.mp.	789
76.	leuprorelin acetate.mp.	185
77.	Leuprolide/	2,010
78.	leuprolide.mp.	2,297
79.	Triptorelin Pamoate/	1,098
80.	triptorelin pamoate.mp.	1,100
81.	triptorelin.mp.	1,228
82.	anamorelin.mp.	29
83.	Abiraterone Acetate/	305
84.	abiraterone acetate.mp.	602
85.	bicalutamide.mp.	1,399
86.	enzalutamide.mp.	846
87.	Flutamide/	1,767
88.	flutamide.mp.	2,452
89.	Diethylstilbestrol/	1,744
90.	diethylstilbestrol.mp.	2,553
91.	interleukin.mp.	246,595
92.	mifamurtide.mp.	61
93.	talimogene laherparepvec.mp.	75
94.	avelumab.mp.	49
95.	Nivolumab/	1,155
96.	nivolumab.mp.	1,645
97.	durvalumab.mp.	100
98.	Ipilimumab/	1,354
99.	ipilimumab.mp.	1,937
100.	necitumumab.mp.	52
101.	pembrolizumab.mp.	1,163
102.	atezolizumab.mp.	281
103.	dinutuximab.mp.	25
104.	olaratumab.mp.	36
105.	Bevacizumab/	10,270
106.	bevacizumab.mp.	13,381
107.	ramucirumab.mp.	355
108.	aflibercept.mp.	1,233
109.	aldesleukin.mp.	111
110.	degarelix.mp.	150
111.	Protein Kinase Inhibitors/	37,421
112.	protein kinase inhibitor*.mp.	39,698
113.	tyrosine kinase inhibitor*.mp.	19,290
114.	tyrosine protein kinase inhibitor*.mp.	41
115.	TKI.mp.	4,410
116.	TKIs.mp.	3,897
117.	multireceptor tyrosine kinase inhibitor*.mp.	6
118.	BRAF kinase inhibitor*.mp.	52

Appendix 2.3 Continued

119.	MAPK inhibitor*.mp.	3,426
120.	mitogen-activated protein kinase inhibitor*.mp.	541
121.	MEK inhibitor*.mp.	4,070
122.	mTOR inhibitor*.mp.	4,675
123.	CDK inhibitor*.mp.	3,450
124.	Cyclin dependent kinase inhibitor*.mp.	25,818
125.	“Poly(ADP-ribose) Polymerase Inhibitors"/	2,727
126.	"Poly(ADP-ribose) Polymerase Inhibitors".mp.	2,822
127.	PARP inhibitor*.mp.	2,055
128.	Hedgehog pathway inhibitor*.mp.	180
129.	Interleukins/	12,436
130.	Interleukins.mp.	15,799
131.	Anti-oestrogen*.mp.	408
132.	Anti-estrogen*.mp.	1,685
133.	Progesterone/	21,305
134.	Progesterone.mp.	50,516
135.	Aromatase Inhibitors/	5,042
136.	Aromatase inhibitor*.mp.	6,875
137.	Gonadotropin-Releasing Hormone/	13,746
138.	Gonadotropin-releasing hormone.mp.	16,400
139.	Androgen Antagonists/	8,314
140.	androgen antagonists.mp.	8,333
141.	Anti-androgen*.mp.	1,848
142.	Androgen Receptor Antagonists/	1,096
143.	Androgen receptor antagonists.mp.	1,165
144.	Estrogens/	25,434
145.	estrogen*.mp.	109,215
146.	oestrogen*.mp.	10,504
147.	Anti-gonadotrophin releasing hormone*.mp.	3
148.	VEGF inhibitor*.mp.	676
149.	Vascular endothelial growth factor inhibitor*.mp.	235
150.	Angiogenesis Inhibitors/	22,825
151.	Angiogenesis inhibitor*.mp.	24,103
152.	EGFR inhibitor*.mp.	2,805
153.	Epidermal growth factor inhibitor*.mp.	51
154.	ALK inhibitor*.mp.	761
155.	cMET inhibitor*.mp.	18
156.	Antibodies, Monoclonal/	107,607
157.	Monoclonal antibod*.mp.	91,652
158.	MAB*.mp.	33,975
159.	exp Leukemia/	108,260
160.	leukemia.mp.	150,191
161.	leukaemia.mp.	18,017
162.	exp Lymphoma/	85,842
163.	lymphoma.mp.	130,464
164.	exp Arthritis, Rheumatoid/	56,908
165.	Rheumatoid arthritis.mp.	58,515
166.	exp Inflammatory Bowel Diseases/	50,019
167.	Inflammatory bowel disease*.mp.	34,931
168.	ulcerative colitis.mp.	21,823
169.	Crohn* disease.mp.	31,880
170.	exp Psoriasis/	21,775
171.	Psoriasis.mp.	24,020
172.	Spondylitis, Ankylosing/	7,001
173.	Ankylosing spondylitis.mp.	7,715
174.	(Or/159 – 173) [OR for MAB uses that we are not interested in]	412,207
175.	156 or 157 or 158 [MAB synonyms]	159,349
176.	175 not 174 [MABs but not disease states we are not interested in]	134,547
177.	targeted treatment*.mp.	3,876
178.	targeted technologies.mp.	10
179.	personalized treatment.mp.	1,770

Appendix 2.3 Continued

180.	personalised treatment.mp.	173
181.	targeted therapy.mp.	35,555
182.	targeted therapies.mp.	13,546
183.	Precision Medicine/	13,508
184.	precision medicine.mp.	15,115
185.	molecularly targeted drug*.mp.	201
186.	Molecular Targeted Therapy/	23,285
187.	Molecular targeted therap*.mp.	24,488
188.	Personalised medicine.mp.	666
189.	Personalized medicine.mp.	6,925
190.	Tailored medicine.mp.	53
191.	Genomic medicine.mp.	800
192.	Precision cancer care.mp.	12
193.	Stratified medicine.mp.	194
194.	(molecular adj2 test).mp.	1,141
195.	genomic testing.mp.	390
196.	targeted test*.mp.	254
197.	biomarker test*.mp.	520
198.	molecular test*.mp.	4,224
199.	mutation* test*.mp.	1,635
200.	test* trend*.mp.	118
201.	test* pattern.mp.1	257
202.	genetic profile/	10
203.	genetic profile.mp.	1,537
204.	Genetic Testing/	31,258
205.	genetic testing.mp.	38,351
206.	Pharmacogenomic Testing/	341
207.	Pharmacogenomic testing.mp.	491
208.	Pharmacogenetic testing.mp.	438
209.	molecular profil*.mp.	4,521
210.	molecular diagnostics.mp.	3,158
211.	genotyping.mp.	47,260
212.	Genomic profiling.mp.	1,132
213.	Gene test*.mp.	1,000
214.	Molecular Diagnostic Techniques/	9,673
215.	Molecular diagnostic techniques.mp.	9,937
216.	(EGFR adj2 test*).mp.	502
217.	HER2 test*.mp.	426
218.	ALK test*.mp.	68
219.	BRAF test*.mp.	69
220.	ER test*.mp.	35
221.	PR test*.mp.	63
222.	KIT test*.mp.	81
223.	BRCA test*.mp.	241
224.	RET test*.mp.	18
225.	MET test*.mp.	26
226.	KRAS test*.mp.	121
227.	NRAS test*.mp.	4
228.	PD1 test*.mp.	1
229.	CTLA4 test*.mp.	1
230.	Receipt of.mp.	12,742
231.	treatment with.mp.	316,651
232.	receive*.mp.	670,956
233.	non-use.mp.	1,743
234.	treatment receipt.mp.	98
235.	receipt.mp.	12,748
236.	initiat*.mp.	393,392
237.	treatment utilisation.mp.	38
238.	Treatment utilisation.mp.	526
239.	utili#ation.mp.	129,587
240.	Access.mp.	202,172

Appendix 2.3 Continued

241.	Underutilisation.mp.	128
242.	Underutilisation.mp.	1,314
243.	Treatment barrier*.mp.	313
244.	Non receipt.mp.	66
245.	Non-initiation.mp.	51
246.	Treatment pattern*.mp.	2,690
247.	Utilisation pattern*.mp.	1,684
248.	Utilisation pattern*.mp.	176
249.	Drug Utilisation/	14,097
250.	Drug utilisation.mp.	279
251.	Drug Utilisation.mp.	17,809
252.	Uptake.mp.	219,854
253.	Provision.mp.	48,132
254.	(Or/230 – 253) [Utilisation synonyms OR]	1,805,579
255.	Socioeconomic status.mp.	25,641
256.	exp Socioeconomic Factors/	284,680
257.	Socioeconomic factor*.mp.	107,585
258.	Poverty.mp.	40,752
259.	Poverty areas.mp.	5,013
260.	Social class.mp.	27,739
261.	Social mobility.mp.	747
262.	Index of multiple deprivation.mp.	401
263.	Socioeconomic position.mp.	2,187
264.	Carstairs index.mp.	40
265.	Townsend Index.mp.	73
266.	Area level deprivation.mp.	117
267.	Poverty level.mp.	1,122
268.	Income level.mp.	1,570
269.	(Income adj6 category).mp.	237
270.	Household income.mp.	6,034
271.	(Median adj2 income).mp.	1,254
272.	Education status.mp.	400
273.	Education adj6 demographic).mp.	1,443
274.	(Education adj6 variable).mp.	343
275.	Education level.mp.	9,065
276.	Employment status.mp.	4,675
277.	Employment characteristic*.mp.	162
278.	(Employment adj6 variable).mp.	81
279.	Unemployed.mp.	5,376
280.	Unemploy* adj6 variable).mp.	30
281.	(Disab* adj6 variable).mp.	380
282.	Race ethnicity.mp.	13,358
283.	(Race adj6 variable).mp.	287
284.	(Ethnicity adj6 variable).mp.	169
285.	Smokers/	594
286.	Smoker*.mp.	57,786
287.	Non-Smokers/	20
288.	Non-smoker*.mp.	9,973
289.	Smoking adj2 history).mp.	8,853
290.	health insurance status.mp.	652
291.	socioeconomic.mp.	139,337
292.	exp Medicare/	29,191
293.	Medicare.mp.	42,291
294.	Race.ti,ab.	68,527
295.	Medicaid/	15,589
296.	medicaid.mp.	29,827
297.	exp Insurance Coverage/	15,004
298.	insurance coverage.mp.	15,415
299.	health plan type.mp.	35
300.	public insurance.mp.	950
301.	Medically Uninsured/	6,285

Appendix 2.3 Continued

302.	Medically uninsured.mp.	6,300
303.	insurance adj6 variable).mp.	92
304.	Commercial insurance.mp.	490
305.	Private insurance.mp.	2,878
306.	Military insurance.mp.	11
307.	Other insurance.mp.	154
308.	Racial.ti,ab.	26,166
309.	Socioeconomics.mp.	363
310.	(Or/1 – 155) or (Or/176 - 193) [Targeted therapy OR]	729,966
311.	(Or/194 – 299) [Molecular Testing OR]	109,759
312.	Or/255 – 309) [Socio-economic status OR]	509,341
313.	310 or 311 [Targeted therapy or molecular testing]	829,983
314.	313 and 254 and 312 [Targeted therapy or molecular testing - and utilisation and socio-economic status]	2,370
315.	limit 314 to yr="1998 -Current"	2,308

B. Embase search strategy 04/02/2019 (kept updated through to 31/12/2019)

	Search Term	**Number Retrieved**
1.	alectinib/	1,139
2.	Alectinib.mp.	1,186
3.	ceritinib/	1,309
4.	Ceritinib.mp.	1,371
5.	crizotinib/	6,436
6.	Crizotinib.mp.	6,748
7.	brigatinib/	437
8.	Brigatinib.mp.	446
9.	erlotinib/	24,988
10	Erlotinib.mp.	25,749
11.	gefitinib/	22,545
12.	Gefitinib.mp.	23,272
13.	afatinib/	4,087
14.	Afatinib.mp.	4,227
15.	osimertinib/	1,822
16.	Osimertinib.mp.	1,892
17.	imatinib/	39,035
18.	Imatinib.mp.	40,550
19.	lapatinib/	10,951
20.	Lapatinib.mp.	11,245
21.	neratinib/	1,254
22.	Neratinib.mp.	1,319
23.	masitinib/	438
24.	Masitinib.mp.	466
25.	lenvatinib/	1,355
26.	Lenvatinib.mp.	1,400
27.	cabozantinib/	2,562
28.	Cabozantinib.mp.	2,670
29.	sunitinib/	20,593
30.	Sunitinib.mp.	21,125
31.	axitinib/	4,067
32.	Axitinib.mp.	4,153
33.	pazopanib/	6,722
34.	Pazopanib.mp.	6,879
35.	tivozanib/	474
36.	Tivozanib.mp.	502
37.	vandetanib/	4,213
38.	Vandetanib.mp.	4,341
39.	nintedanib/	2,126
40.	Nintedanib.mp.	2,208

Appendix 2.3 Continued

41.	regorafenib/	2,744
42.	Regorafenib.mp.	2,847
43.	sorafenib/	25,721
44.	Sorafenib.mp.	26,523
45.	dabrafenib/	3,268
46.	Dabrafenib.mp.	3,433
47.	vemurafenib/	6,445
48.	Vemurafenib.mp.	6,710
49.	encorafenib/	295
50.	Encorafenib.mp.	304
51.	cobimetinib/	844
52.	Cobimetinib.mp.	874
53.	trametinib/	3,718
54.	Trametinib.mp.	3,810
55.	binimetinib/	552
56.	Binimetinib.mp.	559
57.	everolimus/	24,301
58.	Everolimus.mp.	26,840
59.	temsirolimus/	7,606
60.	Temsirolimus.mp.	7,766
61.	abemaciclib/	568
62.	Abemaciclib.mp.	580
63.	palbociclib/	2,151
64.	Palbociclib.mp.	2,198
65.	ribociclib/	708
66.	Ribociclib.mp.	737
67.	niraparib/	617
68.	Niraparib.mp.	633
69.	olaparib/	3,493
70.	Olaparib.mp.	3,603
71.	rucaparib/	801
72.	Rucaparib.mp.	814
73.	vismodegib/	1,705
74.	Vismodegib.mp.	1,775
75.	sonidegib/	589
76.	Sonidegib.mp.	604
77.	cetuximab/	25,558
78.	Cetuximab.mp.	26,316
79.	panitumumab/	7,398
80.	Panitumumab.mp.	7,683
81.	pertuzumab/	3,931
82.	Pertuzumab.mp.	4,083
83.	trastuzumab/	34,473
84.	Trastuzumab.mp.	36,566
85.	trastuzumab emtansine/	2,113
86.	Trastuzumab emtansine.mp.	2,189
87.	tamoxifen/	49,834
88.	Tamoxifen.mp.	53,451
89.	fulvestrant/	7,848
90.	Fulvestrant.mp.	8,037
91.	toremifene/	1,823
92.	Toremifene.mp.	1,881
93.	megestrol acetate/	3,681
94.	Megestrol acetate.mp.	3,811
95.	medroxyprogesterone acetate/	11,339
96.	Medroxyprogesterone acetate.mp.	12,949
97.	anastrozole/	9,033
98.	Anastrozole.mp.	9,210
99.	exemestane/	5,710
100.	Exemestane.mp.	5,874
101.	letrozole/	10,862

Appendix 2.3 Continued

102.	Letrozole.mp.	11,108
103.	goserelin/	5,609
104.	Goserelin.mp.	5,681
105.	buserelin/	2,424
106.	Buserelin.mp.	3,087
107.	leuprorelin/	9,131
108.	Leuprorelin.mp.	9,167
109.	Leuprolide.mp.	2,203
110.	triptorelin/	4,254
111.	Triptorelin.mp.	4,319
112.	anamorelin/	183
113.	Anamorelin.mp.	194
114.	abiraterone/	3,327
115.	Abiraterone.mp.	5,572
116.	abiraterone acetate/	2,330
117.	Abiraterone Acetate.mp.	2,548
118.	bicalutamide/	5,589
119.	Bicalutamide.mp.	5,728
120.	enzalutamide/	4,338
121.	Enzalutamide.mp.	4,561
122.	flutamide/	6,280
123.	Flutamide.mp.	6,535
124.	diethylstilbestrol/	5,394
125.	Diethylstilbestrol.mp.	5,888
126.	interleukin.mp.	564,845
127.	mifamurtide/	707
128.	Mifamurtide.mp.	722
129.	talimogene laherparepvec/	626
130.	Talimogene laherparepvec.mp.	651
131.	avelumab/	1,135
132.	Avelumab.mp.	1,164
133.	nivolumab/	10,264
134.	Nivolumab.mp.	10,680
135.	durvalumab/	1,896
136.	Durvalumab.mp.	1,930
137.	ipilimumab/	10,466
138.	Ipilimumab.mp.	10,820
139.	necitumumab/	355
140.	Necitumumab.mp.	367
141.	pembrolizumab/	8,373
142.	Pembrolizumab.mp.	8,723
143.	atezolizumab/	2,548
144.	Atezolizumab.mp.	2,635
145.	dinutuximab/	230
146.	Dinutuximab.mp.	238
147.	olaratumab/	249
148.	Olaratumab.mp.	258
149.	bevacizumab/	51,869
150.	Bevacizumab.mp.	53,282
151.	ramucirumab/	1,940
152.	Ramucirumab.mp.	2,302
153.	aflibercept/	4,787
154.	Aflibercept.mp.	4,928
155.	aldesleukin/	212
156.	Aldesleukin.mp.	431
157.	degarelix/	687
158.	Degarelix.mp.	723
159.	protein kinase inhibitor/	9,667
160.	Protein kinase inhibitor*.mp.	25,072
161.	protein tyrosine kinase inhibitor/	28,485
162.	Protein tyrosine kinase inhibitor*.mp.	28,885

Appendix 2.3 Continued

163.	Tyrosine kinase inhibitor*.mp.	50,599
164.	TKI.mp.	15,510
165.	TKIs.mp.	11,350
166.	Multireceptor tyrosine kinase inhibitor*.mp.	13
167.	B Raf kinase inhibitor/	1,595
168.	B Raf kinase inhibitor*.mp.	1,615
169.	MAPK inhibitor*.mp.	4,758
170.	mitogen activated protein kinase inhibitor/	8,696
171.	Mitogen activated protein kinase inhibitor*.mp.	9,120
172.	MEK inhibitor*.mp.	7,524
173.	mammalian target of rapamycin inhibitor"/	10,191
174.	"Mammalian target of rapamycin inhibitor*".mp.	10,944
175.	mTOR inhibitor*.mp.	11,349
176.	CDK inhibitor*.mp.	4,688
177.	cyclin dependent kinase inhibitor/	6,489
178.	Cyclin dependent kinase inhibitor*.mp.	30,048
179.	nicotinamide adenine dinucleotide adenosine diphosphate ribosyltransferase inhibitor/	5,101
180.	nicotinamide adenine dinucleotide adenosine diphosphate ribosyltransferase inhibitor*.mp.	5,103
181.	"Poly(ADP-ribose) polymerase inhibitor*".mp.	837
182.	PARP inhibitor*.mp.	5,062
183.	Hedgehog pathway inhibitor*.mp.	442
184.	interleukin derivative/	4,093
185.	Interleukins.mp.	7,979
186.	Anti-oestrogen*.mp.	602
187.	antiestrogen*/	6,857
188.	antiestrogen.mp.	8,332
189.	progesterone/	52,758
190.	Progesterone*.mp.	93,935
191.	aromatase inhibitor/	12,922
192.	Aromatase inhibitor*.mp.	16,109
193.	Gonadotropin-releasing hormone.mp.	11,285
194.	Androgen antagonist*.mp.	425
195.	antiandrogen/	9,750
196.	Antiandrogen*.mp.	12,668
197.	androgen receptor antagonist/	804
198.	Androgen receptor antagonist*.mp.	1,280
199.	estrogen/	77,835
200.	Estrogen*.mp.	193,483
201.	Oestrogen*.mp.	15,780
202.	Anti-gonadotrophin releasing hormone*.mp.	4
203.	VEGF inhibitor*.mp.	1,475
204.	Vascular endothelial growth factor inhibitor*.mp.	424
205.	angiogenesis inhibitor/	16,797
206.	Angiogenesis inhibitor*.mp.	18,898
207.	epidermal growth factor receptor kinase inhibitor/	7,515
208.	Epidermal growth factor receptor kinase inhibitor*.mp.	7,555
209.	EGFR inhibitor*.mp.	5,966
210.	anaplastic lymphoma kinase inhibitor/	722
211.	anaplastic lymphoma kinase inhibitor*.mp.	815
212.	ALK inhibitor*.mp.	1,986
213.	cMET inhibitor*.mp.	77
214.	monoclonal antibody/	128,259
215.	monoclonal antibod*.mp.	193,938
216.	MAB*.mp.	52,680
217.	exp leukemia/	203,033
218.	Leukemia.mp.	308,280
219.	Leukaemia.mp.	30,062
220.	exp lymphoma/	209,181
221.	Lymphoma.mp.	231,956

Appendix 2.3 Continued

222.	exp rheumatoid arthritis/	143,804
223.	Rheumatoid arthritis.mp.	156,410
224.	exp inflammatory bowel disease/	113,514
225.	Inflammatory bowel disease*.mp.	72,223
226.	Ulcerative colitis.mp.	60,175
227.	Crohn* disease.mp.	77,747
228.	exp psoriasis/	63,401
229.	Psoriasis.mp.	57,412
230.	exp ankylosing spondylitis/	19,351
231.	Ankylosing spondylitis.mp.	22,848
232.	(Or/217 - 231) [OR for MAB uses that we are not interested in]	837,019
233.	214 or 215 or 216[MAB synonyms]	211,516
234.	234. 233 not 232 [MABs but not disease states we are not interested in]	178,024
235.	Targeted treatment*.mp.	8,327
236.	Targeted technologies.mp.	21
237.	Personalized treatment.mp.	4,160
238.	Personalised treatment.mp.	481
239.	Targeted therapy.mp.	54,912
240.	Targeted therapies.mp.	29,215
241.	Precision medicine.mp.	8,630
242.	Molecularly targeted drug*.mp.	395
243.	molecularly targeted therapy/	27,560
244.	Molecular targeted therap*.mp.	3,761
245.	personalized medicine/	33,557
246.	Personalized medicine.mp.	39,734
247.	Personalised medicine.mp.	1,727
248.	Tailored medicine.mp.	104
249.	Genomic medicine.mp.	1,445
250.	Precision cancer care.mp.	31
251.	Stratified medicine.mp.	437
252.	(Or/1 – 213) or (Or/234 – 251) [Targeted Therapy OR]	1,412,394
253.	(Molecular adj2 test).mp.	2,508
254.	Genomic testing.mp.	966
255.	Targeted test*.mp.	501
256.	biomarker test*.mp.	1,401
257.	molecular test*.mp.	9,599
258.	mutation* test*.mp.	4,099
259.	Test* trend*.mp.	268
260.	Genetic profile.mp.	3,145
261.	Genetic testing.mp.	28,563
262.	pharmacogenetic testing/	623
263.	Pharmacogenetic testing.mp.	1,319
264.	Pharmacogenomic testing.mp.	353
265.	Molecular profil*.mp.	10,439
266.	Molecular diagnostics.mp.	7,419
267.	Genotyping.mp.	79,715
268.	Genomic profiling.mp.	3,262
269.	Gene test*.mp.	2,238
270.	Molecular diagnostic techniques.mp.	598
271.	(EGFR adj2 test*).mp.	1,972
272.	HER2 test*.mp.	962
273.	ALK test*.mp.	274
274.	BRAF test*.mp.	240
275.	ER test*.mp.	81
276.	PR test*.mp.	140
277.	KIT test*.mp.	168
278.	BRCA test*.mp.	621
279.	RET test*.mp.	34
280.	MET test*.mp.	55
281.	KRAS test*.mp.	363
282.	NRAS test*.mp.	18

Appendix 2.3 Continued

283.	PD1 test*.mp.	1
284.	CTLA4 test*.mp.	1
285.	(Or/253 – 284) [Molecular Testing OR]	150,822
286.	Receipt of.mp.	22,756
287.	Treatment with.mp.	554,115
288.	Receive*.mp.	1,315,723
289.	Non-use.mp.	3,195
290.	Treatment receipt.mp.	200
291.	Receipt.mp.	22,766
292.	Initiat*.mp.	704,192
293.	Treatment utilisation.mp.	77
294.	Utili#ation.mp.	290,352
295.	Access.mp.	400,446
296.	Underutilisation.mp.	254
297.	Underutilisation.mp.	2,326
298.	Treatment barrier*.mp.	628
299.	Non receipt.mp.	152
300.	Non-initiation.mp.	113
301.	Treatment pattern*.mp.	7,916
302.	Utilisation pattern*.mp.	3,461
303.	Utilisation pattern*.mp.	325
304.	drug utilisation/	16,809
305.	Drug utilisation.mp.	18,761
306.	Drug utilisation.mp.	884
307.	Uptake.mp.	360,750
308.	Provision.mp.	78,138
309.	(Or/286 – 308) [Utilisation OR]	3,314,589
310.	Socioeconomic status.mp.	39,767
311.	Socioeconomic factor*.mp.	9,588
312.	poverty/	35,806
313.	Poverty.mp.	44,341
314.	Poverty areas.mp.	196
315.	social class/	21,008
316.	Social class.mp.	23,178
317.	Social mobility.mp.	511
318.	Index of multiple deprivation.mp.	911
319.	Socioeconomic position.mp.	2,926
320.	Carstairs index.mp.	60
321.	Townsend index.mp.	101
322.	Area level deprivation.mp.	170
323.	Poverty level.mp.	1,859
324.	socioeconomics/	101,475
325.	Socioeconomic.mp.	86,230
326.	Socioeconomics.mp.	101,733
327.	Income level.mp.	2,729
328.	(Income adj6 category).mp.	419
329.	household income/	3,729
330.	Household income.mp.	11,616
331.	Median adj2 income).mp.	2,792
332.	Education status.mp.	976
333.	(Education adj6 demographic).mp.	2,862
334.	(Education adj6 variable).mp.	1,062
335.	Education level.mp.	17,048
336.	Employment status.mp.	15,256
337.	Employment characteristic*.mp.	206
338.	Employment adj6 variable).mp.	182
339.	Unemployed.mp.	9,462
340.	Unemploy* adj6 variable).mp.	48
341.	Disab* adj6 variable).mp.	906
342.	Race ethnicity.mp.	24,035
343.	(Race adj6 variable).mp.	672

Appendix 2.3 Continued

344.	Ethnicity adj6 variable).mp.	435
345.	Race.ti,ab.	125,659
346.	Racial.ti,ab.	40,988
347.	Smoker*.mp.	104,222
348.	Non-smoker*.mp.	19,714
349.	(Smoking adj2 history).mp.	21,857
350.	Health insurance status.mp.	1,049
351.	exp medicare/	60,158
352.	Medicare.mp.	70,152
353.	medicaid/	34,820
354.	Medicaid.mp.	41,596
355.	Insurance coverage.mp.	8,902
356.	Health plan type.mp.	111
357.	Public insurance.mp.	1,805
358.	medically uninsured/	4,169
359.	Medically uninsured.mp.	4,286
360.	Insurance adj6 variable).mp.	263
361.	Commercial insurance.mp.	1,357
362.	private health insurance/	4,141
363.	Private health insurance.mp.	5,222
364.	Military insurance.mp.	35
365.	Other insurance.mp.	314
366.	(Or/310 – 365) [Socio-economic Status]	580,384
367.	252 or 285 [Targeted therapy or molecular testing]	1,541,395
368.	367 and 309 and 366 Targeted therapy or molecular testing – utilisation and Socio-economic status]	7,203
369.	limit 368 to yr="1998 -current"	7,137

C. Scopus search strategy 05/02/2019 (kept updated through to 31/12/2019)

	Search Term	**Number Retrieved**
1.	Alectinib" OR "Ceritinib" OR "Crizotinib" OR "Brigatinib" OR "Erlotinib" OR "Gefitinib" OR "Afatinib" OR "Osimertinib" OR "Imatinib" OR "Lapatinib" OR "Neratinib" OR "Masitinib" OR "Lenvatinib" OR "Cabozantinib" OR "Sunitinib" OR "Axitinib"	78,221
2.	"Pazopanib" OR "Tivozanib" OR "Vandetanib" or "Nintedanib" OR "Regorafenib" OR "Sorafenib" OR "Dabrafenib" OR "Vemurafenib" OR "Encorafenib" OR "Cobimetinib" OR "Trametinib" OR "Binimetinib" OR "Everolimus" OR "Temsirolimus" OR "Abemaciclib"	47,541
3.	Palbociclib" OR "Ribociclib" OR "Niraparib" OR "Olaparib" OR "Rucaparib" OR "Vismodegib" OR "Sonidegib" OR "Cetuximab" OR "Pantimumab" OR "Pertuzumab" OR "Trastuzumab" OR "Trastuzumab Emtansine" OR "Tamoxifen" OR "Fulvestrant" OR "Toremifene"	106,598
4.	"Megestrol Acetate" OR "Medroxyprogesterone Acetate" OR "Anastrozole" OR "Exemestane" OR "Letrozole" OR "Goserelin" OR "Buserelin" OR "Leuprorelin Acetate" OR "Leuprolide" OR "Triptorelin Pamoate" OR "Triptorelin" OR "Anamorelin" OR "Abiraterone Acetate"	52,197
5.	"Bicalutamide" OR "Enzalutamide" OR "Flutamide" OR "Diethylstilbestrol" OR "Interleukin" OR "Mifamurtide" OR "Talimogene Laherparepvec" OR "Avelumab" OR "Nivolumab" OR "Durvalumab" OR "Ipilimumab" OR "Necitumumab" OR "Pembrolizumab" OR "Atezolizumab"	565,325
6.	"Dinutuximab" OR "Olaratumab" OR "Bevacizumab" OR "Ramucirumab" OR "Aflibercept" OR "Aldesleukin" OR "Degarelix"	45,136
7.	"Protein Kinase Inhibitor*" OR "Tyrosine Kinase Inhibitor*" OR "Tyrosine Protein Kinase Inhibitor*" OR "TKI" or "TKIs" OR "Multireceptor Tyrosine Kinase Inhibitor*" OR "BRAF Kinase Inhibitor*" OR "MAPK Inhibitor"	86,043
8.	"Mitogen-activated Protein Kinase Inhibitor*" OR "mTOR Inhibitor*" OR "CDK Inhibitor*" OR "Cyclin Dependent Kinase Inhibitor*" OR "Poly(ADP-ribose) Polymerase Inhibitor*" OR "PARP Inhibitor*" OR "Hedgehog Pathway Inhibitor*" OR "Interleukins"	76,259

Appendix 2.3 Continued

9.	"Anti-oestrogen*" OR "Progesterone" OR "Aromatase Inhibitor*" OR "Gonadotrophin-Releasing Hormone*" OR "Androgen Antagonist*" OR "Anti-androgen*" OR "Androgen Receptor Antagonist*" OR "Estrogen*" OR "Oestrogen*" OR "Anti-gonadotrophin Releasing Hormone*"	375,362
10	"VEGF Inhibitor*" OR "Vascular Endothelial Growth Factor Inhibitor*" OR "Angiogenesis Inhibitor*" OR "EGFR Inhibitor*" OR "Epidermal Growth Factor Inhibitor*" OR "ALK Inhibitor*" OR "cMET Inhibitor*"	36,188
11.	"Monoclonal Antibod*" OR "MAB"	307,997
12.	"Leukemia" OR "Leukaemia" OR "Lymphoma" OR "Rheumatoid Arthritis" OR "Inflammatory Bowel Diseases" OR "Ulcerative Colitis" OR "Crohn* Disease" OR "Psoriasis" OR "Ankylosing Spondylitis"	982,757
13.	("Monoclonal Antibod*" OR "MAB") AND NOT (#12) [MABs but not disease states that we are not interested in]	265,890
14.	"Targeted Treatment*" OR "Targeted Technologies" OR "Personalised Treatment" OR "Personalized Treatment" OR "Targeted Therapy" OR "Targeted Therapies" OR "Precision Medicine" OR "Molecularly Targeted Drug*" OR "Molecularly Targeted Therap*"	87,157
15.	"Personalised Medicine" OR "Personalized Medicine" OR "Tailored Medicine" OR "Genomic Medicine" OR "Precision Cancer Care" OR "Stratified Medicine"	37,306
16.	(#1) OR (#2) OR (#3) OR (#4) OR (#5) OR (#6) OR (#7) OR (#8) OR (#9) OR (#10) OR (#13) OR (#14) OR (#15) [Targeted Therapies OR]	1,496,121
17.	"Genomic Testing" OR "Targeted Test*" OR "Biomarker Test*" OR "Molecular Test*" OR "Mutation* Test*" OR "Test* Trend*" OR "Test* Pattern" OR "Genetic Profile" OR "Genetic Testing" OR "Pharmacogenomic Testing*" OR "Pharmacogenetic Testing"	58,536
18.	"Molecular Profil*" OR "Molecular Diagnostics" OR "Genotyping" OR "Genomic Profiling" OR "Gene Test*" OR "Molecular Diagnostic Techniques" OR "EGFR Test*" OR "HER2 Test*" OR "ALK Test*" OR "BRAF Test*" OR "ER Test*" OR "PR Test*" OR "KIT Test*"	99,486
19.	"BRCA Test*" OR "RET Test*" OR "MET Test*" OR "KRAS Test*" OR "NRAS Test*" OR "PD1 Test*" OR "CTLA4 Test*"	597
20.	(#17) OR (#18) OR (#19) [Molecular Testing OR]	153,771
21.	"Receipt of" OR "Treatment with" OR "Receive*" OR "Non-use" OR "Treatment Receipt" OR "Receipt" OR "Initiat*" OR "Treatment Utilisation" OR "Treatment Utilisation" OR "Utilisation" OR "Utilisation" OR "Access" OR "Underutilisation" OR "Underutilisation"	5,132,936
22.	"Treatment Barrier*" OR "Non-receipt" OR "Non-initiation" OR "Treatment Pattern*" OR "Utilisation Pattern*" OR "Utilisation Pattern*" OR "Drug Utilisation" OR "Drug Utilisation" OR "Uptake" OR "Provision"	848,630
23.	(#21) OR (#22) [Utilisation synonyms]	5,816,262
24.	"Socioeconomic Status" OR "Socioeconomic Factors" OR "Poverty" OR "Poverty Areas" OR "Social Class" OR "Social Mobility" OR "Index of Multiple Deprivation" OR "Socioeconomic Position" OR "Carstairs Index" OR "Townsend Index" OR "Area Level Deprivation"	335,186
25.	"Poverty Level" OR "Income Level" Or "Household Income" "Education Status" OR "Education Level" OR "Employment Status" OR "Employment Characteristic*" OR "Unemployed" OR "Race Ethnicity" OR "Smoker*" OR "Non-smoker*" OR "Smoking History"	3,630
26.	"Health Insurance Status" OR "Socioeconomic*" OR "Medicare" OR "Medicaid" OR "Insurance Coverage" OR "Health Plan Type" OR "Public Insurance" OR "Medically Uninsured" OR "Commercial Insurance" OR "Private Insurance" OR "Military Insurance"	415,616
27.	"Other Insurance" OR "Income Category" OR "Median Income" OR "Education Demographic" OR "Education Variable" OR "Employment Variable" OR "Unemploy* Variable" OR "Disab* Variable" OR "Race Variable" OR "Ethnicity Variable" OR "Insurance Variable"	2,670
28.	(#24) OR (#25) OR (#26) OR (#27) [Socio-economic status OR]	554,623
29.	(#16) OR (#20) [Targeted therapy or molecular testing OR]	1,633,613
30.	(#29) AND (#23) AND (#28) [Targeted therapy or molecular testing – and utilisation socio-economic status] Search filtered for date: 1998 onwards.	1,916

Appendix 2.3 Continued
D. CINAHL search strategy 06/02/2019 (kept updated through to 31/12/2019)

	Search Term	Number Retrieved
1.	TX "Alectinib" OR "Ceritinib" OR "Crizotinib" OR "Brigatinib" OR "Erlotinib" OR "Gefitinib" OR "Afatinib" OR "Osimertinib" OR "Imatinib" OR "Lapatinib" OR "Neratinib" OR "Masitinib" OR "Lenvatinib" OR "Cabozantinib" OR "Sunitinib" OR "Axitinib" OR "Pazopanib" OR "Tivozanib" OR "Vandetanib" or "Nintedanib" OR "Regorafenib" OR "Sorafenib" OR "Dabrafenib" OR "Vemurafenib" OR "Encorafenib" OR "Cobimetinib" OR "Trametinib" OR "Binimetinib" OR "Everolimus" OR "Temsirolimus" OR "Abemaciclib" OR "Palbociclib" OR "Ribociclib" OR "Niraparib" OR "Olaparib" OR "Rucaparib" OR "Vismodegib" OR "Sonidegib" OR "Cetuximab" OR "Pantimumab" OR "Pertuzumab" OR "Trastuzumab" OR "Trastuzumab Emtansine" OR "Tamoxifen" OR "Fulvestrant" OR "Toremifene" OR "Megestrol Acetate" OR "Medroxyprogesterone Acetate" OR "Anastrozole" OR "Exemestane" OR "Letrozole" OR "Goserelin" OR "Buserelin" OR "Leuprorelin Acetate" OR "Leuprolide" OR "Triptorelin Pamoate" OR "Triptorelin" OR "Anamorelin" OR "Abiraterone Acetate" OR "Bicalutamide" OR "Enzalutamide" OR "Flutamide" OR "Diethylstilbestrol" OR "Interleukin" OR "Mifamurtide" OR "Talimogene Laherparepvec" OR "Avelumab" OR "Nivolumab" OR "Durvalumab" OR "Ipilimumab" OR "Necitumumab" OR "Pembrolizumab" OR "Atezolizumab" OR "Dinutuximab" OR "Olaratumab" OR "Bevacizumab" OR "Ramucirumab" OR "Aflibercept" OR "Aldesleukin" OR "Degarelix"	42,832
2.	TX "Protein Kinase Inhibitor*" OR "Tyrosine Kinase Inhibitor*" OR "Tyrosine Protein Kinase Inhibitor*" OR "TKI" or "TKIs" OR "Multireceptor Tyrosine Kinase Inhibitor*" OR "BRAF Kinase Inhibitor*" OR "MAPK Inhibitor" OR "Mitogen-activated Protein Kinase Inhibitor*" OR "mTOR Inhibitor*" OR "CDK Inhibitor*" OR "Cyclin Dependent Kinase Inhibitor*" OR "Poly(ADP-ribose) Polymerase Inhibitor*" OR "PARP Inhibitor*" OR "Hedgehog Pathway Inhibitor*" OR "Interleukins" OR "Anti-oestrogen*" OR "Progesterone" OR "Aromatase Inhibitor*" OR "Gonadotrophin-Releasing Hormone*" OR "Androgen Antagonist*" OR "Anti-androgen*" OR "Androgen Receptor Antagonist*" OR "Estrogen*" OR "Oestrogen*" OR "Anti-gonadotrophin Releasing Hormone*" OR "VEGF Inhibitor*" OR "Vascular Endothelial Growth Factor Inhibitor*" OR "Angiogenesis Inhibitor*" OR "EGFR Inhibitor*" OR "Epidermal Growth Factor Inhibitor*" OR "ALK Inhibitor*" OR "cMET Inhibitor*"	55,984
3.	TX ("Monoclonal Antibod*" OR "MAB") NOT ("Leukemia" OR "Leukaemia" OR "Lymphoma" OR "Rheumatoid Arthritis" OR "Inflammatory Bowel Diseases" OR "Ulcerative Colitis" OR "Crohn* Disease" OR "Psoriasis" OR "Ankylosing Spondylitis") [MABs but not disease states that we are not interested in]	6,393
4.	TX "Targeted Treatment*" OR "Targeted Technologies" OR "Personalised Treatment" OR "Personalized Treatment" OR "Targeted Therapy" OR "Targeted Therapies" OR "Precision Medicine" OR "Molecularly Targeted Drug*" OR "Molecularly Targeted Therap*" OR "Personalised Medicine" OR "Personalized Medicine" OR "Tailored Medicine" OR "Genomic Medicine" OR "Precision Cancer Care" OR "Stratified Medicine"	13,902
5.	S1 OR S2 OR S3 OR S4 [Targeted Therapies OR]	100,005

Appendix 2.3 Continued

6.	TX “Genomic Testing” OR “Targeted Test*” OR “Biomarker Test*” OR “Molecular Test*” OR “Mutation* Test*” OR “Test* Trend*” OR “Test* Pattern” OR “Genetic Profile” OR “Genetic Testing” OR “Pharmacogenomic Testing*” OR “Pharmacogenetic Testing” OR “Molecular Profil*” OR “Molecular Diagnostics” OR “Genotyping” OR “Genomic Profiling” OR “Gene Test*” OR “Molecular Diagnostic Techniques” OR “EGFR Test*” OR “HER2 Test*” OR “ALK Test*” OR “BRAF Test*” OR “ER Test*” OR “PR Test*” OR “KIT Test*” OR ““BRCA Test*” OR “RET Test*” OR “MET Test*” OR “KRAS Test*” OR “NRAS Test*” OR “PD1 Test*” OR “CTLA4 Test*” [Molecular Testing OR]	17,044
7.	TX “Receipt of” OR “Treatment with” OR “Receive*” OR “Non-use” OR “Treatment Receipt” OR “Receipt” OR “Initiat*” OR “Treatment Utilisation” OR “Treatment Utilisation” OR “Utilisation” OR “Utilisation” OR “Access” OR “Underutilisation” OR “Underutilisation” OR “Treatment Barrier*” OR “Non-receipt” OR “Non-initiation” OR “Treatment Pattern*” OR “Utilisation Pattern*” OR “Utilisation Pattern*” OR “Drug Utilisation” OR “Drug Utilisation” OR “Uptake” OR “Provision” [Utilisation OR]	1,384,830
8.	TX “Socioeconomic Status” OR “Socioeconomic Factors” OR “Poverty” OR “Poverty Areas” OR “Social Class” OR “Social Mobility” OR “Index of Multiple Deprivation” OR “Socioeconomic Position” OR “Carstairs Index” OR “Townsend Index” OR “Area Level Deprivation” OR “Poverty Level” OR “Income Level” Or “Household Income” “Education Status” OR “Education Level” OR “Employment Status” OR “Employment Characteristic*” OR “Unemployed” OR “Race Ethnicity” OR “Smoker*” OR “Non-smoker*” OR “Smoking History” OR ““Health Insurance Status” OR “Socioeconomic*” OR “Medicare” OR “Medicaid” OR “Insurance Coverage” OR “Health Plan Type” OR “Public Insurance” OR “Medically Uninsured” OR “Commercial Insurance” OR “Private Insurance” OR “Military Insurance” OR “Other Insurance” OR “Income Category” OR “Median Income” OR “Education Demographic” OR “Education Variable” OR “Employment Variable” OR “Unemploy* Variable” OR “Disab* Variable” OR “Race Variable” OR “Ethnicity Variable” OR “Insurance Variable” [Socio-economic Status OR]	239,101
9.	S5 or S6 [Targeted Therapies OR Molecular Testing]	115,100
10.	S7 AND S8 AND S9 [Utilisation AND Socio-economic Status AND Targeted Therapies OR Molecular Testing]	1,273
11	Search filtered for date: 1998 onwards	1,246

Appendix 2.3 Continued

E. Web of Science search strategy 06/02/2019 (kept updated through to 31/12/2019)

	Search Term	Number Retrieved
1.	TS=("Alectinib" OR "Ceritinib" OR "Crizotinib" OR "Brigatinib" OR "Erlotinib" OR "Gefitinib" OR "Afatinib" OR "Osimertinib" OR "Imatinib" OR "Lapatinib" OR "Neratinib" OR "Masitinib" OR "Lenvatinib" OR "Cabozantinib" OR "Sunitinib" OR "Axitinib" OR "Pazopanib" OR "Tivozanib" OR "Vandetanib" or "Nintedanib" OR "Regorafenib" OR "Sorafenib" OR "Dabrafenib" OR "Vemurafenib" OR "Encorafenib" OR "Cobimetinib" OR "Trametinib" OR "Binimetinib" OR "Everolimus" OR "Temsirolimus" OR "Abemaciclib" OR "Palbociclib" OR "Ribociclib" OR "Niraparib" OR "Olaparib" OR "Rucaparib" OR "Vismodegib" OR "Sonidegib" OR "Cetuximab" OR "Pantimumab" OR "Pertuzumab" OR "Trastuzumab" OR "Trastuzumab Emtansine" OR "Tamoxifen" OR "Fulvestrant" OR "Toremifene" OR "Megestrol Acetate" OR "Medroxyprogesterone Acetate" OR "Anastrozole" OR "Exemestane" OR "Letrozole" OR "Goserelin" OR "Buserelin" OR "Leuprorelin Acetate" OR "Leuprolide" OR "Triptorelin Pamoate" OR "Triptorelin" OR "Anamorelin" OR "Abiraterone Acetate" OR "Bicalutamide" OR "Enzalutamide" OR "Flutamide" OR "Diethylstilbestrol" OR "Interleukin" OR "Mifamurtide" OR "Talimogene Laherparepvec" OR "Avelumab" OR "Nivolumab" OR "Durvalumab" OR "Ipilimumab" OR "Necitumumab" OR "Pembrolizumab" OR "Atezolizumab" OR "Dinutuximab" OR "Olaratumab" OR "Bevacizumab" OR "Ramucirumab" OR "Aflibercept" OR "Aldesleukin" OR "Degarelix")	406,270
2.	TS=("Protein Kinase Inhibitor*" OR "Tyrosine Kinase Inhibitor*" OR "Tyrosine Protein Kinase Inhibitor*" OR "TKI" or "TKIs" OR "Multireceptor Tyrosine Kinase Inhibitor*" OR "BRAF Kinase Inhibitor*" OR "MAPK Inhibitor" OR "Mitogen-activated Protein Kinase Inhibitor*" OR "mTOR Inhibitor*" OR "CDK Inhibitor*" OR "Cyclin Dependent Kinase Inhibitor*" OR "Poly(ADP-ribose) Polymerase Inhibitor*" OR "PARP Inhibitor*" OR "Hedgehog Pathway Inhibitor*" OR "Interleukins" OR "Anti-oestrogen*" OR "Progesterone" OR "Aromatase Inhibitor*" OR "Gonadotrophin-Releasing Hormone*" OR "Androgen Antagonist*" OR "Anti-androgen*" OR "Androgen Receptor Antagonist*" OR "Estrogen*" OR "Oestrogen*" OR "Anti-gonadotrophin Releasing Hormone*" OR "VEGF Inhibitor*" OR "Vascular Endothelial Growth Factor Inhibitor*" OR "Angiogenesis Inhibitor*" OR "EGFR Inhibitor*" OR "Epidermal Growth Factor Inhibitor*" OR "ALK Inhibitor*" OR "cMET Inhibitor*")	258,973
3.	TS=("Monoclonal Antibod*" OR "MAB")	157,839
4.	TS=("Leukemia" OR "Leukaemia" OR "Lymphoma" OR "Rheumatoid Arthritis" OR "Inflammatory Bowel Diseases" OR "Ulcerative Colitis" OR "Crohn* Disease" OR "Psoriasis" OR "Ankylosing Spondylitis")	596,216
5.	#3 NOT #4 [MABs but not disease states that we are not interested in]	134,963
6.	TS=("Targeted Treatment*" OR "Targeted Technologies" OR "Personalised Treatment" OR "Personalized Treatment" OR "Targeted Therapy" OR "Targeted Therapies" OR "Precision Medicine" OR "Molecularly Targeted Drug*" OR "Molecularly Targeted Therap*" OR "Personalised Medicine" OR "Personalized Medicine" OR "Tailored Medicine" OR "Genomic Medicine" OR "Precision Cancer Care" OR "Stratified Medicine")	64,813
7.	#1 OR #2 OR #5 OR #6 [Targeted Therapies OR]	771,432
8.	TS=("Genomic Testing" OR "Targeted Test*" OR "Biomarker Test*" OR "Molecular Test*" OR "Mutation* Test*" OR "Test* Trend*" OR "Test* Pattern" OR "Genetic Profile" OR "Genetic Testing" OR "Pharmacogenomic Testing*" OR "Pharmacogenetic Testing" OR "Molecular Profil*" OR "Molecular Diagnostics" OR "Genotyping" OR "Genomic Profiling" OR "Gene Test*" OR "Molecular Diagnostic Techniques" OR "EGFR Test*" OR "HER2 Test*" OR "ALK Test*" OR "BRAF Test*" OR "ER Test*" OR "PR Test*" OR "KIT Test*" OR ""BRCA Test*" OR "RET Test*" OR "MET Test*" OR "KRAS Test*" OR "NRAS Test*" OR "PD1 Test*" OR "CTLA4 Test*") [Molecular Testing OR]	101,713

Appendix 2.3 Continued

9.	TS=("Receipt of" OR "Treatment with" OR "Receive*" OR "Non-use" OR "Treatment Receipt" OR "Receipt" OR "Initiat*" OR "Treatment Utilisation" OR "Treatment Utilisation" OR "Utilisation" OR "Utilisation" OR "Access" OR "Underutilisation" OR "Underutilisation" OR "Treatment Barrier*" OR "Non-receipt" OR "Non-initiation" OR "Treatment Pattern*" OR "Utilisation Pattern*" OR "Utilisation Pattern*" OR "Drug Utilisation" OR "Drug Utilisation" OR "Uptake" OR "Provision") [Utilisation OR]	3,190,637
10.	TS=("Socioeconomic Status" OR "Socioeconomic Factors" OR "Poverty" OR "Poverty Areas" OR "Social Class" OR "Social Mobility" OR "Index of Multiple Deprivation" OR "Socioeconomic Position" OR "Carstairs Index" OR "Townsend Index" OR "Area Level Deprivation" OR "Poverty Level" OR "Income Level" Or "Household Income" "Education Status" OR "Education Level" OR "Employment Status" OR "Employment Characteristic*" OR "Unemployed" OR "Race Ethnicity" OR "Smoker*" OR "Non-smoker*" OR "Smoking History" OR ""Health Insurance Status" OR "Socioeconomic*" OR "Medicare" OR "Medicaid" OR "Insurance Coverage" OR "Health Plan Type" OR "Public Insurance" OR "Medically Uninsured" OR "Commercial Insurance" OR "Private Insurance" OR "Military Insurance" OR "Other Insurance" OR "Income Category" OR "Median Income" OR "Education Demographic" OR "Education Variable" OR "Employment Variable" OR "Unemploy* Variable" OR "Disab* Variable" OR "Race Variable" OR "Ethnicity Variable" OR "Insurance Variable") [Socio-economic Status OR]	330,771
11.	#7 or #8 [Targeted Therapies OR Molecular Testing]	863,528
12.	#11 AND #9 AND #10 [Targeted Therapies OR Molecular Testing AND Utilisation AND Socio-economic Status] Timespan: 1998-2019	1,730

F. PubMed search strategy 06/02/2019 (kept updated through to 31/12/2019)

	Search Term	**Number Retrieved**
1.	Search (Alectinib" OR "Ceritinib" OR "Crizotinib" OR "Brigatinib" OR "Erlotinib" OR "Gefitinib" OR "Afatinib" OR "Osimertinib" OR "Imatinib" OR "Lapatinib" OR "Neratinib" OR "Masitinib" OR "Lenvatinib" OR "Cabozantinib" OR "Sunitinib" OR "Axitinib" OR "Pazopanib" OR "Tivozanib" OR "Vandetanib" or "Nintedanib" OR "Regorafenib" OR "Sorafenib" OR "Dabrafenib" OR "Vemurafenib" OR "Encorafenib" OR "Cobimetinib" OR "Trametinib" OR "Binimetinib" OR "Everolimus" OR "Temsirolimus" OR "Abemaciclib" OR "Palbociclib" OR "Ribociclib" OR "Niraparib" OR "Olaparib" OR "Rucaparib" OR "Vismodegib" OR "Sonidegib" OR "Cetuximab" OR "Pantimumab" OR "Pertuzumab" OR "Trastuzumab" OR "Trastuzumab Emtansine" OR "Tamoxifen" OR "Fulvestrant" OR "Toremifene" OR "Megestrol Acetate" OR "Medroxyprogesterone Acetate" OR "Anastrozole" OR "Exemestane" OR "Letrozole" OR "Goserelin" OR "Buserelin" OR "Leuprorelin Acetate" OR "Leuprolide" OR "Triptorelin Pamoate" OR "Triptorelin" OR "Anamorelin" OR "Abiraterone Acetate" OR "Bicalutamide" OR "Enzalutamide" OR "Flutamide" OR "Diethylstilbestrol" OR "Interleukin" OR "Mifamurtide" OR "Talimogene Laherparepvec" OR "Avelumab" OR "Nivolumab" OR "Durvalumab" OR "Ipilimumab" OR "Necitumumab" OR "Pembrolizumab" OR "Atezolizumab" OR "Dinutuximab" OR "Olaratumab" OR "Bevacizumab" OR "Ramucirumab" OR "Aflibercept" OR "Aldesleukin" OR "Degarelix")	466,277

Appendix 2.3 Continued

2.	Search (“Protein Kinase Inhibitor*” OR “Tyrosine Kinase Inhibitor*” OR “Tyrosine Protein Kinase Inhibitor*” OR “TKI” or “TKIs” OR “Multireceptor Tyrosine Kinase Inhibitor*” OR “BRAF Kinase Inhibitor*” OR “MAPK Inhibitor” OR “Mitogen-activated Protein Kinase Inhibitor*” OR “mTOR Inhibitor*” OR “CDK Inhibitor*” OR “Cyclin Dependent Kinase Inhibitor*” OR “Poly(ADP-ribose) Polymerase Inhibitor*” OR “PARP Inhibitor*” OR “Hedgehog Pathway Inhibitor*” OR “Interleukins” OR “Anti-oestrogen*” OR “Progesterone” OR “Aromatase Inhibitor*” OR “Gonadotrophin-Releasing Hormone*” OR “Androgen Antagonist*” OR “Anti-androgen*” OR “Androgen Receptor Antagonist*” OR “Estrogen*” OR “Oestrogen*” OR “Anti-gonadotrophin Releasing Hormone*” OR “VEGF Inhibitor*” OR “Vascular Endothelial Growth Factor Inhibitor*” OR “Angiogenesis Inhibitor*” OR “EGFR Inhibitor*” OR “Epidermal Growth Factor Inhibitor*” OR “ALK Inhibitor*” OR “cMET Inhibitor*”)	306,142
3.	Search (“Monoclonal Antibod*” OR “MAB”)	41,708
4.	Search (“Leukemia” OR “Leukaemia” OR “Lymphoma” OR “Rheumatoid Arthritis” OR “Inflammatory Bowel Diseases” OR “Ulcerative Colitis” OR “Crohn* Disease” OR “Psoriasis” OR “Ankylosing Spondylitis”)	730,455
5.	Search (#3 NOT #4) [MABs but not disease states that we are not interested in]	38,509
6.	Search (“Targeted Treatment*” OR “Targeted Technologies” OR “Personalised Treatment” OR “Personalized Treatment” OR “Targeted Therapy” OR “Targeted Therapies” OR “Precision Medicine” OR “Molecularly Targeted Drug*” OR “Molecularly Targeted Therap*” OR “Personalised Medicine” OR “Personalized Medicine” OR “Tailored Medicine” OR “Genomic Medicine” OR “Precision Cancer Care” OR “Stratified Medicine”)	96,878
7.	Search (#1 OR #2 OR #5 OR #6) [Targeted Therapies OR]	824,944
8.	Search (“Genomic Testing” OR “Targeted Test*” OR “Biomarker Test*” OR “Molecular Test*” OR “Mutation* Test*” OR “Test* Trend*” OR “Test* Pattern” OR “Genetic Profile” OR “Genetic Testing” OR “Pharmacogenomic Testing*” OR “Pharmacogenetic Testing” OR “Molecular Profil*” OR “Molecular Diagnostics” OR “Genotyping” OR “Genomic Profiling” OR “Gene Test*” OR “Molecular Diagnostic Techniques” OR “EGFR Test*” OR “HER2 Test*” OR “ALK Test*” OR “BRAF Test*” OR “ER Test*” OR “PR Test*” OR “KIT Test*” OR ““BRCA Test*” OR “RET Test*” OR “MET Test*” OR “KRAS Test*” OR “NRAS Test*” OR “PD1 Test*” OR “CTLA4 Test*”) [Molecular Testing OR]	205,023
9.	Search (“Receipt of” OR “Treatment with” OR “Receive*” OR “Non-use” OR “Treatment Receipt” OR “Receipt” OR “Initiat*” OR “Treatment Utilisation” OR “Treatment Utilisation” OR “Utilisation” OR “Utilisation” OR “Access” OR “Underutilisation” OR “Underutilisation” OR “Treatment Barrier*” OR “Non-receipt” OR “Non-initiation” OR “Treatment Pattern*” OR “Utilisation Pattern*” OR “Utilisation Pattern*” OR “Drug Utilisation” OR “Drug Utilisation” OR “Uptake” OR “Provision”) [Utilisation OR]	1,552,648
10.	Search (“Socioeconomic Status” OR “Socioeconomic Factors” OR “Poverty” OR “Poverty Areas” OR “Social Class” OR “Social Mobility” OR “Index of Multiple Deprivation” OR “Socioeconomic Position” OR “Carstairs Index” OR “Townsend Index” OR “Area Level Deprivation” OR “Poverty Level” OR “Income Level” Or “Household Income” “Education Status” OR “Education Level” OR “Employment Status” OR “Employment Characteristic*” OR “Unemployed” OR “Race Ethnicity” OR “Smoker*” OR “Non-smoker*” OR “Smoking History” OR ““Health Insurance Status” OR “Socioeconomic*” OR “Medicare” OR “Medicaid” OR “Insurance Coverage” OR “Health Plan Type” OR “Public Insurance” OR “Medically Uninsured” OR “Commercial Insurance” OR “Private Insurance” OR “Military Insurance” OR “Other Insurance” OR “Income Category” OR “Median Income” OR “Education Demographic” OR “Education Variable” OR “Employment Variable” OR “Unemploy* Variable” OR “Disab* Variable” OR “Race Variable” OR “Ethnicity Variable” OR “Insurance Variable”) [Socio-economic Status OR]	485,253

Appendix 2.3 Continued

11.	Search (#7 or #8) [Targeted Therapies OR Molecular Testing]	1,016,885
12.	Search (#11 AND #9 AND #10) [Targeted Therapies OR Molecular Testing AND Utilisation AND Socio-economic Status] Search filtered for date: 1998 onwards	2,424

G. PsycInfo search strategy 07/02/2019 (kept updated through to 31/12/2019)

	Search Term	Number Retrieved
1.	alectinib.mp	2
2.	ceritinib.mp.	2
3.	Crizotinib/	0
4.	crizotinib.mp.	6
5.	brigatinib.mp.	0
6.	Erlotinib Hydrochloride/	0
7.	erlotinib.mp.	27
8.	Gefitinib/	0
9.	gefitinib.mp.	16
10.	Afatinib/	0
11.	afatinib.mp.	5
12.	osimertinib.mp.	1
13.	Imatinib Mesylate/	0
14.	imatinib.mp.	50
15.	Lapatinib/	0
16.	lapatinib.mp.	8
17.	neratinib.mp.	1
18.	masitinib.mp.	6
19.	lenvatinib.mp.	0
20.	cabozantinib.mp.	1
21.	Sunitinib/	0
22.	sunitinib.mp.	37
23.	Axitinib/	0
24.	axitinib.mp.	3
25.	pazopanib.mp.	6
26.	tivozanib.mp.	0
27.	vandetanib.mp.	5
28.	nintedanib.mp.	3
29.	regorafenib.mp.	0
30.	Sorafenib/	0
31.	sorafenib.mp.	17
32.	dabrafenib.mp.	7
33.	Vemurafenib/	0
34.	vemurafenib.mp.	10
35.	encorafenib.mp.	0
36.	cobimetinib.mp.	0
37.	trametinib.mp.	5
38.	binimetinib.mp.	0
39.	Everolimus/	0
40.	everolimus.mp.	59
41.	temsirolimus.mp.	15
42.	abemaciclib.mp.	1
43.	palbociclib.mp.	2
44.	ribociclib.mp.	1
45.	niraparib.mp.	1
46.	olaparib.mp.	3
47.	rucaparib.mp.	4
48.	vismodegib.mp.	3
49.	sonidegib.mp.	0
50.	Cetuximab/	0
51.	cetuximab.mp.	12

Appendix 2.3 Continued

52.	Panitumumab/	0
53.	panitumumab.mp.	4
54.	pertuzumab.mp.	0
55.	Trastuzumab/	0
56.	trastuzumab.mp.	32
57.	trastuzumab emtansine.mp.	0
58.	exp Tamoxifen/	0
59.	tamoxifen.mp.	520
60.	Fulvestrant/	0
61.	fulvestrant.mp.	20
62.	toremifene.mp.	4
63.	Megestrol Acetate/	0
64.	megestrol acetate.mp.	20
65.	Medroxyprogesterone Acetate/	0
66.	medroxyprogesterone acetate.mp.	207
67.	Anastrozole/	0
68.	anastrozole.mp.	31
69.	exemestane.mp.	14
70.	Letrozole/	0
71.	letrozole.mp.	73
72.	Goserelin/	0
73.	goserelin.mp.	30
74.	Buserelin/	0
75.	buserelin.mp.	6
76.	leuprorelin acetate.mp.	6
77.	Leuprolide/	0
78.	leuprolide.mp.	83
79.	Triptorelin Pamoate/	0
80.	triptorelin pamoate.mp.	0
81.	triptorelin.mp.	28
82.	anamorelin.mp.	0
83.	Abiraterone Acetate/	0
84.	abiraterone acetate.mp.	2
85.	bicalutamide.mp.	9
86.	enzalutamide.mp.	3
87.	Flutamide/	0
88.	flutamide.mp.	146
89.	Diethylstilbestrol/	0
90.	diethylstilbestrol.mp.	89
91.	interleukin.mp.	6,496
92.	mifamurtide.mp.	0
93.	talimogene laherparepvec.mp.	1
94.	avelumab.mp.	1
95.	Nivolumab/	0
96.	nivolumab.mp.	15
97.	durvalumab.mp.	0
98.	Ipilimumab/	0
99.	ipilimumab.mp.	16
100.	necitumumab.mp.	0
101.	pembrolizumab.mp.	6
102.	atezolizumab.mp.	1
103.	dinutuximab.mp.	2
104.	olaratumab.mp.	0
105.	Bevacizumab/	0
106.	bevacizumab.mp.	117
107.	ramucirumab.mp.	0
108.	aflibercept.mp.	5
109.	aldesleukin.mp.	0
110.	degarelix.mp.	4
111.	Protein Kinase Inhibitors/	0
112.	protein kinase inhibitor*.mp.	139

Appendix 2.3 Continued

113.	tyrosine kinase inhibitor*.mp.	207
114.	tyrosine protein kinase inhibitor*.mp.	1
115.	TKI.mp.	41
116.	TKIs.mp.	24
117.	multireceptor tyrosine kinase inhibitor*.mp.	0
118.	BRAF kinase inhibitor*.mp.	0
119.	MAPK inhibitor*.mp.	149
120.	mitogen-activated protein kinase inhibitor*.mp.	28
121.	MEK inhibitor*.mp.	177
122.	mTOR inhibitor*.mp.	147
123.	CDK inhibitor*.mp.	55
124.	Cyclin dependent kinase inhibitor*.mp.	78
125.	“Poly(ADP-ribose) Polymerase Inhibitors"/	0
126.	"Poly(ADP-ribose) Polymerase Inhibitors".mp.	4
127.	PARP inhibitor*.mp.	42
128.	Hedgehog pathway inhibitor*.mp.	1
129.	Interleukins/	3,030
130.	Interleukins.mp.	3,168
131.	Anti-oestrogen*.mp.	6
132.	Anti-estrogen*.mp.	36
133.	Progesterone/	1,599
134.	Progesterone.mp.	3,326
135.	Aromatase Inhibitors/	0
136.	Aromatase inhibitor*.mp.	239
137.	Gonadotropin-Releasing Hormone/	0
138.	Gonadotropin-releasing hormone.mp.	716
139.	Androgen Antagonists/	176
140.	androgen antagonists.mp.	12
141.	Anti-androgen*.mp.	91
142.	Androgen Receptor Antagonists/	0
143.	Androgen receptor antagonists.mp.	30
144.	Estrogens/	3,274
145.	estrogen*.mp.	6,726
146.	oestrogen*.mp.	726
147.	Anti-gonadotrophin releasing hormone*.mp.	0
148.	VEGF inhibitor*.mp.	8
149.	Vascular endothelial growth factor inhibitor*.mp.	2
150.	Angiogenesis Inhibitors/	0
151.	Angiogenesis inhibitor*.mp.	31
152.	EGFR inhibitor*.mp.	17
153.	Epidermal growth factor inhibitor*.mp.	2
154.	ALK inhibitor*.mp.	5
155.	cMET inhibitor*.mp.	0
156.	Antibodies, Monoclonal/	0
157.	Monoclonal antibod*.mp.	1,289
158.	MAB*.mp.	785
159.	exp Leukemia/	1,055
160.	leukemia.mp.	1,794
161.	leukaemia.mp.	256
162.	exp Lymphoma/	0
163.	lymphoma.mp.	1,190
164.	exp Arthritis, Rheumatoid/	1,655
165.	Rheumatoid arthritis.mp.	2,484
166.	exp Inflammatory Bowel Diseases/	0
167.	Inflammatory bowel disease*.mp.	823
168.	ulcerative colitis.mp.	410
169.	Crohn* disease.mp.	537
170.	exp Psoriasis/	0
171.	Psoriasis.mp.	496
172.	Spondylitis, Ankylosing/	0
173.	Ankylosing spondylitis.mp.	135

Appendix 2.3 Continued

174.	(Or/159 – 173) [OR for MAB uses that we are not interested in]	7,152
175.	156 or 157 or 158 [MAB synonyms]	1,891
176.	175 not 174 [MABs but not disease states we are not interested in]	1,838
177.	targeted treatment*.mp.	599
178.	targeted technologies.mp.	1
179.	personalized treatment.mp.	314
180.	personalised treatment.mp.	29
181.	targeted therapy.mp.	183
182.	targeted therapies.mp.	280
183.	Precision Medicine/	0
184.	precision medicine.mp.	325
185.	molecularly targeted drug*.mp.	0
186.	Molecular Targeted Therapy/	0
187.	Molecular targeted therap*.mp.	13
188.	Personalised medicine.mp.	61
189.	Personalized medicine.mp.	683
190.	Tailored medicine.mp.	2
191.	Genomic medicine.mp.	93
192.	Precision cancer care.mp.	0
193.	Stratified medicine.mp.	22
194.	(molecular adj2 test).mp.	22
195.	genomic testing.mp.	80
196.	targeted test*.mp.	55
197.	biomarker test*.mp.	55
198.	molecular test*.mp.	107
199.	mutation* test*.mp.	45
200.	test* trend*.mp.	29
201.	test* pattern.mp.1	127
202.	genetic profile/	0
203.	genetic profile.mp.	138
204.	Genetic Testing/	1,640
205.	genetic testing.mp.	2,996
206.	Pharmacogenomic Testing/	0
207.	Pharmacogenomic testing.mp.	29
208.	Pharmacogenetic testing.mp.	63
209.	molecular profil*.mp.	162
210.	molecular diagnostics.mp.	57
211.	genotyping.mp.	2,492
212.	Genomic profiling.mp.	25
213.	Gene test*.mp.	108
214.	Molecular Diagnostic Techniques/	0
215.	Molecular diagnostic techniques.mp.	5
216.	(EGFR adj2 test*).mp.	2
217.	HER2 test*.mp.	2
218.	ALK test*.mp.	0
219.	BRAF test*.mp.	1
220.	ER test*.mp.	9
221.	PR test*.mp.	30
222.	KIT test*.mp.	5
223.	BRCA test*.mp.	33
224.	RET test*.mp.	3
225.	MET test*.mp.	4
226.	KRAS test*.mp.	0
227.	NRAS test*.mp.	0
228.	PD1 test*.mp.	0
229.	CTLA4 test*.mp.	0
230.	Receipt of.mp.	6,250
231.	treatment with.mp.	35,679
232.	receive*.mp.	174,402
233.	non-use.mp.	786
234.	treatment receipt.mp.	72

Appendix 2.3 Continued

235.	receipt.mp.	6,550
236.	initiat*.mp.	94,430
237.	treatment utilisation.mp.	13
238.	Treatment utilisation.mp.	669
239.	utili#ation.mp.	37,196
240.	Access.mp.	91,960
241.	Underutilisation.mp.	34
242.	Underutilisation.mp.	828
243.	Treatment barrier*.mp.	4,529
244.	Non receipt.mp.	23
245.	Non-initiation.mp.	17
246.	Treatment pattern*.mp.	443
247.	Utilisation pattern*.mp.	660
248.	Utilisation pattern*.mp.	33
249.	Drug Utilisation/	0
250.	Drug utilisation.mp.	25
251.	Drug Utilisation.mp.	271
252.	Uptake.mp.	13,696
253.	Provision.mp.	31,784
254.	(Or/230 – 253) [Utilisation synonyms OR]	440,273
255.	Socioeconomic status.mp.	32,313
256.	exp Socioeconomic Factors/	0
257.	Socioeconomic factor*.mp.	2,272
258.	Poverty.mp.	21,065
259.	Poverty areas.mp.	535
260.	Social class.mp.	9,017
261.	Social mobility.mp.	1,832
262.	Index of multiple deprivation.mp.	285
263.	Socioeconomic position.mp.	989
264.	Carstairs index.mp.	12
265.	Townsend Index.mp.	21
266.	Area level deprivation.mp.	43
267.	Poverty level.mp.	642
268.	Income level.mp.	13,101
269.	(Income adj6 category).mp.	111
270.	Household income.mp.	3,267
271.	(Median adj2 income).mp.	469
272.	Education status.mp.	317
273.	Education adj6 demographic).mp.	2,036
274.	(Education adj6 variable).mp.	476
275.	Education level.mp.	6,045
276.	Employment status.mp.	16,168
277.	Employment characteristic*.mp.	187
278.	(Employment adj6 variable).mp.	127
279.	Unemployed.mp.	5,487
280.	Unemploy* adj6 variable).mp.	32
281.	(Disab* adj6 variable).mp.	211
282.	Race ethnicity.mp.	9,927
283.	(Race adj6 variable).mp.	272
284.	(Ethnicity adj6 variable).mp.	241
285.	Smokers/	0
286.	Smoker*.mp.	19,565
287.	Non-Smokers/	0
288.	Non-smoker*.mp.	1,688
289.	Smoking adj2 history).mp.	1,415
290.	health insurance status.mp.	270
291.	socioeconomic.mp.	46,048
292.	exp Medicare/	1,893
293.	Medicare.mp.	4,633
294.	Race.ti,ab.	51,272
295.	Medicaid/	2,088

Appendix 2.3 Continued

296.	medicaid.mp.	5,406
297.	exp Insurance Coverage/	0
298.	insurance coverage.mp.	1,982
299.	health plan type.mp.	11
300.	public insurance.mp.	318
301.	Medically Uninsured/	350
302.	Medically uninsured.mp.	30
303.	insurance adj6 variable).mp.	31
304.	Commercial insurance.mp.	91
305.	Private insurance.mp.	764
306.	Military insurance.mp.	5
307.	Other insurance.mp.	40
308.	Racial.ti,ab.	35,306
309.	Socioeconomics.mp.	283
310.	(Or/1 – 155) or (Or/176 - 193) [Targeted therapy OR]	22,891
311.	(Or/194 – 299) [Molecular Testing OR]	6,397
312.	Or/255 – 309) [Socio-economic status OR]	197,552
313.	310 or 311 [Targeted therapy or molecular testing]	29,013
314.	313 and 254 and 312 [Targeted therapy or molecular testing - and utilisation and socio-economic status]	172
315.	limit 314 to yr="1998 -Current"	169

Appendix 2.4 Author Initials for Chapter 2

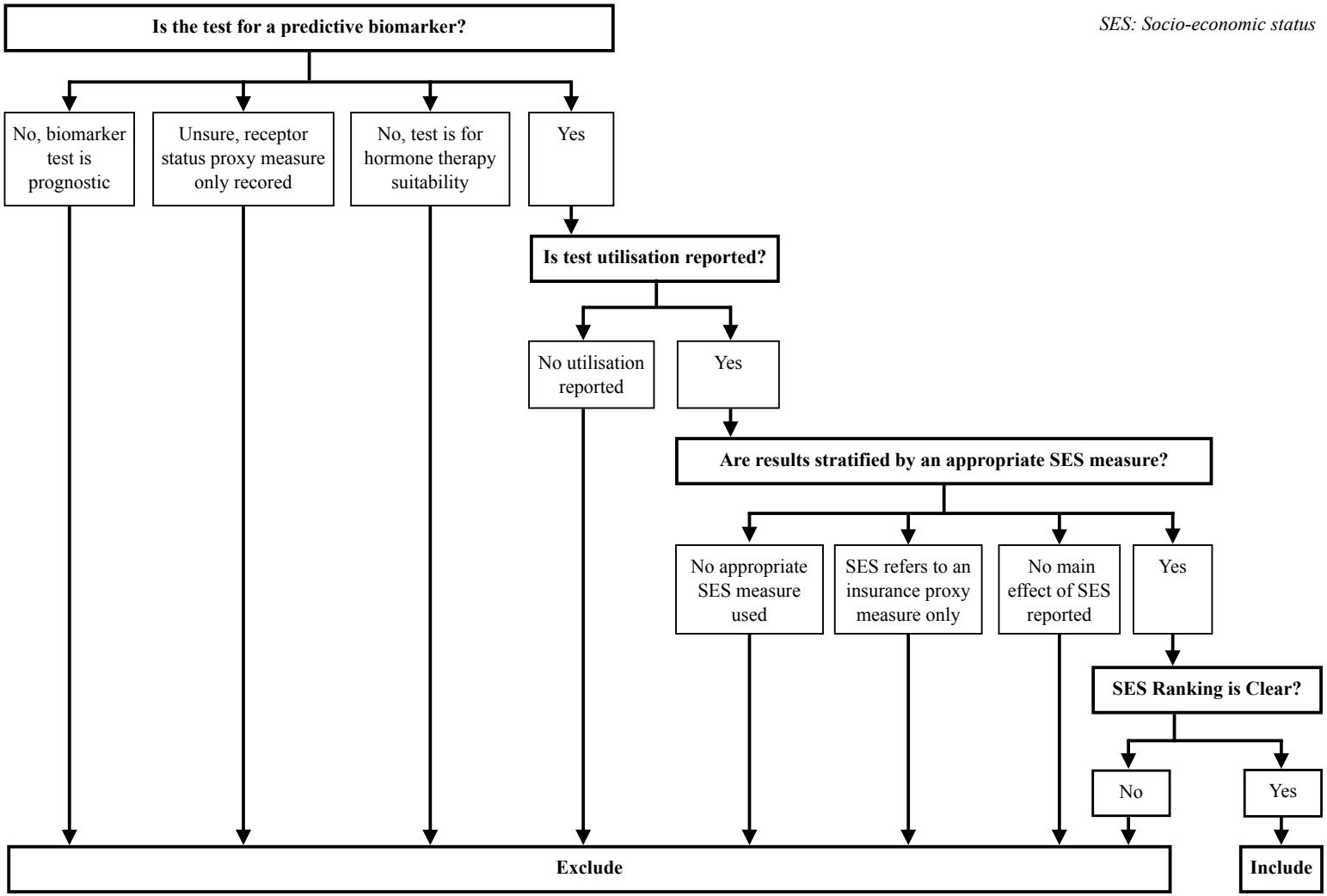
Author initials as per the published review paper:

Norris. R. P, Dew. R, Sharp. L, Greystoke. A, Rice. S, Johnell. K and Todd. A (2020) Are there socio-economic inequalities in utilisation of predictive biomarker tests and biological and precision therapies for cancer? A systematic review and meta-analysis, BMC Medicine 18, 282, https://doi.org/10.1186/s12916-020-01753-0. (231) (Appendix 2.1)

AG: Dr. Alastair Greystoke
AT: Professor Adam Todd
KJ: Professor Kristina Johnell
LS: Professor Linda Sharp
RN: Ruth P Norris
SR: Stephen Rice

Appendix 2.5 Full text study inclusion/exclusion criteria decision trees for Chapter 2

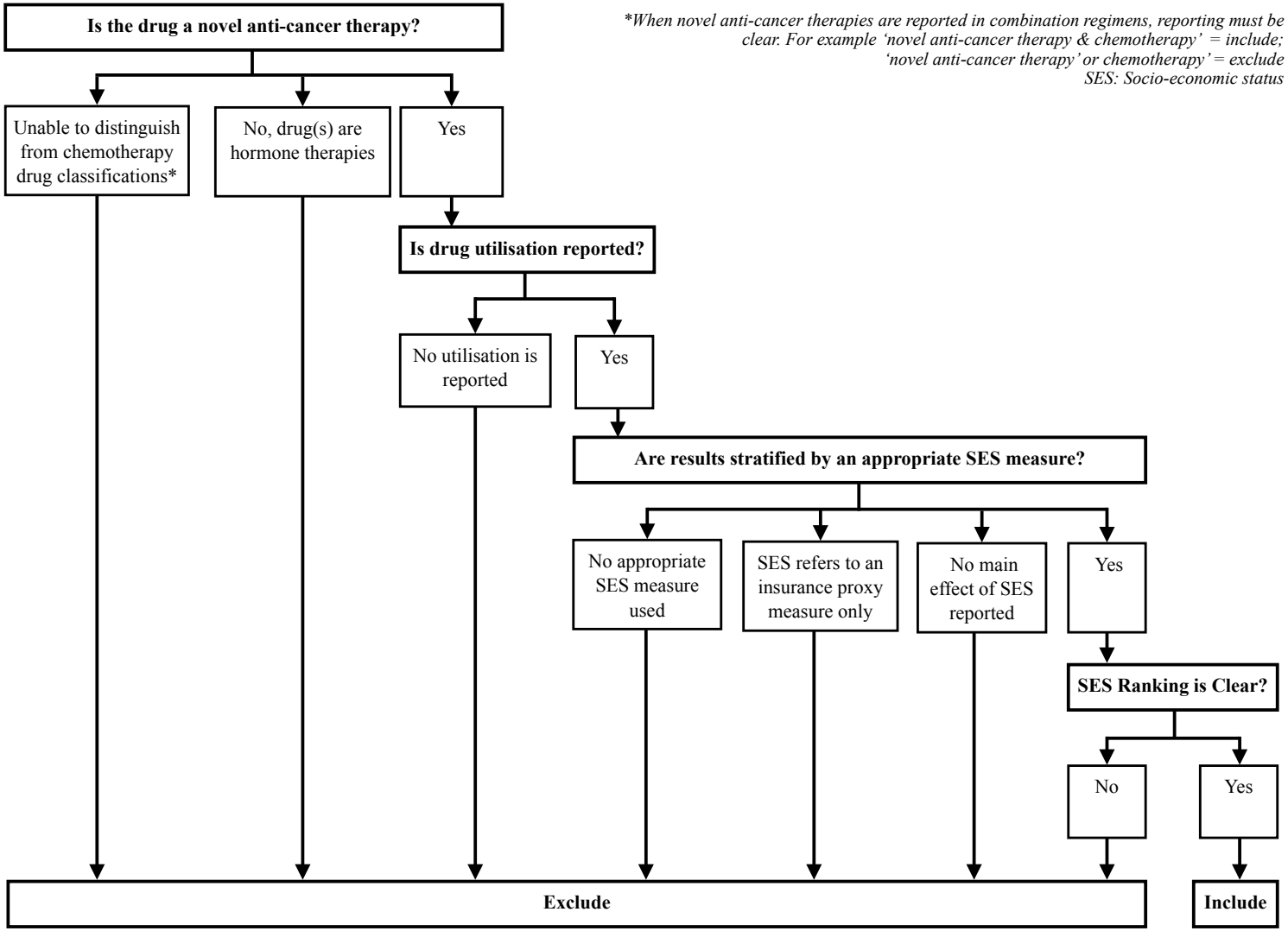
A. Predictive biomarker test articles



SES: Socio-economic status

Appendix 2.5 Continued

B. Novel anti-cancer therapy articles



When novel anti-cancer therapies are reported in combination regimens, reporting must be clear. For example 'novel anti-cancer therapy & chemotherapy' = include; 'novel anti-cancer therapy' or chemotherapy' = exclude

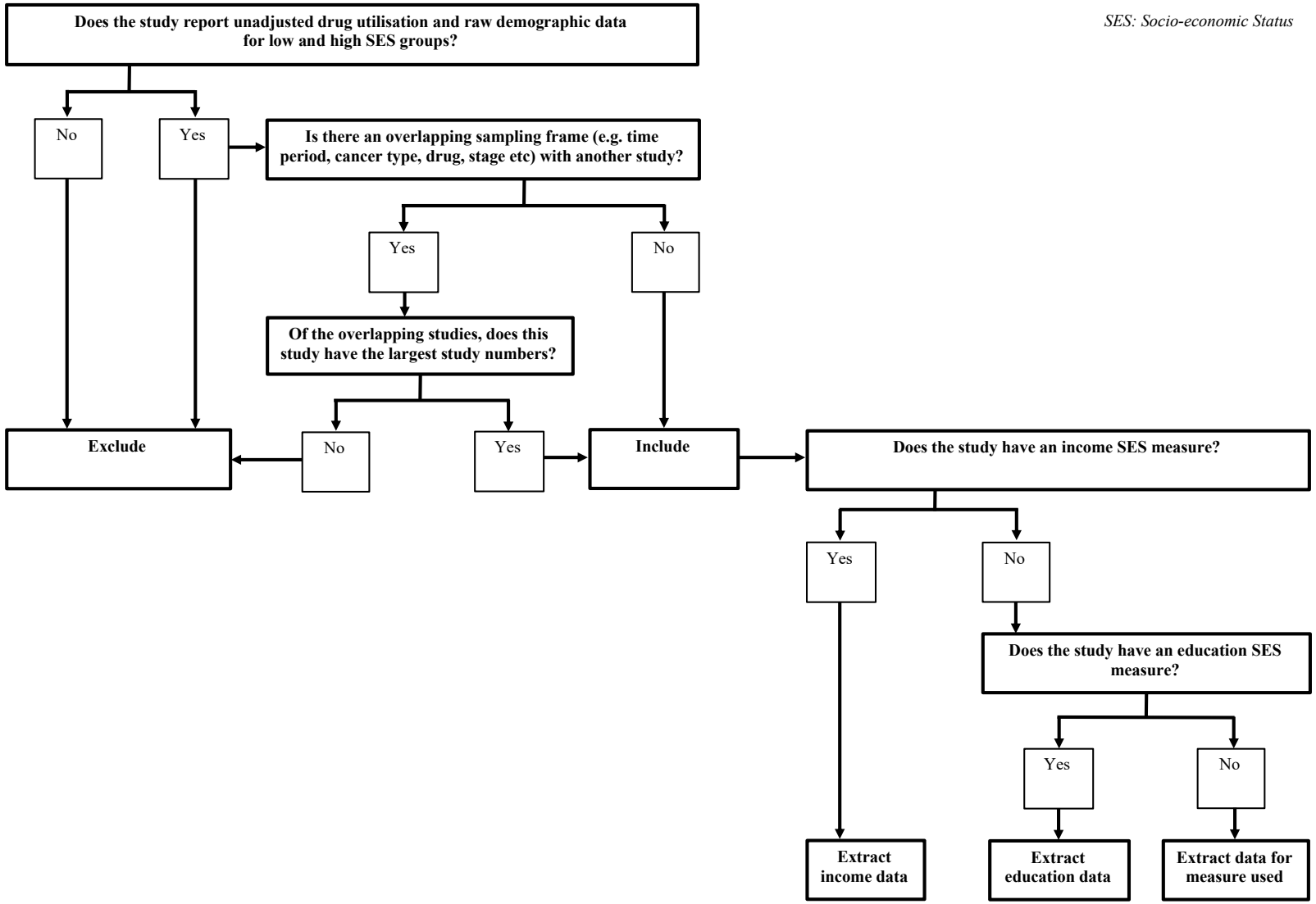
SES: Socio-economic status

Appendix 2.6 Quality Appraisal Tool for Chapter 2

Question	Appraisal	Yes = 1	No = 0	Unclear = 0
	Data Sources			
1	Did the author(s) address issues regarding completeness of SES and treatment data or consistency of coding SES and treatment data in the registry/database? *Missing data are addressed.* *Details of data source given.* *2 points = Tick* *1 point = Half tick*			
	Methods - Study Population & Variables			
2	Was the study subjects and the setting described in detail? *Explicit statement for inclusion/exclusion criteria given e.g. cancer type, time period, staging, location, age.* *Total study population number of interest stated (explicit statement, on consort diagram or total listed in table - not requiring calculation).* *2 points = Tick* *1 point = Half tick*			
	Methods - Operational Definitions			
3	Was the exposure (SES) measure clear? *Unit of measure is stated.* *Clear which SES rank is high or low.* *2 points = Tick* *1 point = Half tick*			
4	Was the outcome measure (drug or test utilisation) clear? *Comparator(s) reported e.g. no novel anti-cancer therapy and/or predictive biomarker test or a clinical alternative e.g. chemotherapy.* *Drug code identification in registry listed or enough detail provided to identify where this information came from.* *2 points = Tick* *1 point = Half tick*			
	Results & Statistics			
5	Is utilisation data of interest reported in tables as patient numbers (not just percent)? *Yes = Tick* *No = No tick*			
6	Were the SES and treatment groups that were statistically compared related to data of interest? *P value compares SES differences between the treatment groups of interest (e.g. no novel anti-cancer therapy and/or predictive biomarker test or a clinical alternative e.g. chemotherapy).* *Describes the statistical test used to compare SES difference in treatment groups (e.g. Chi square).* *2 points = Tick* *1 point = Half tick*			
7	Has the association between exposure and outcome being statistically analysed (e.g. OR, RR) and this reported for the variables of interest? Yes = Tick No = No tick			
8	Control methods: Did the authors use a method to control for confounders within the data of interest? *Adjusted analysis is carried out for the data of interest e.g. multivariate analysis, PSM.* *Confounders are listed.* *2 points = Tick* *1 point = Half tick*			
	Discussion/Conclusions			
9	Have the author(s) discussed SES and treatment utilisation findings? *Yes = Tick* *No = No tick*			
10	Have the authors acknowledged limitations that may reduce the generalisability of the results to other populations and settings? *Yes = Tick (Clear - both limitations and generalisability addressed).* *Yes= Half a tick (Ambiguous, more reliant on "assumptions" than explicit statements).* *No = No tick*			
	Overall Score			

OR: Odds Ratio; PSM: Propensity score matching; RR: Risk Ratio; SES: Socio-economic status

Appendix 2.7 Meta-analysis inclusion/exclusion criteria for the 48 narrative review studies decision tree for Chapter 2



Appendix 2.8 Characteristics of studies reporting predictive biomarker test and/or novel anti-cancer therapies utilisation without a denominator population or which only reported a measure of average socio-economic status for Chapter 2

	Sampling Frame					SES		Utilisation by SES Grouping (Number, %)				
Study	**Country**	**Data Source**	**Study Population[a]**	**Predictive Biomarker Test/Biological and Precision Medicine Overall Utilisation (Number, %)**	**Comparator**	**Unit**	**Measure**	**Lowest SES Group**			**Highest SES Group**	**QA**
Breast Cancer: Novel Anti-Cancer Therapies												
Freedman *et al.* (2014) (268)	USA	SEER-Medicare	28% of US Population 2005 - 2009 Age ≥ 66 Stage I - III n = 2,106	Concurrent or Sequential Trastuzumab (with Either Standard or Non-Standard Chemotherapy) n = 2,106 (100.0)	None	Census Tract	Median Household Income	Q1 (Low) 514/2,106 (24.4)	Q2 534/2,106 (25.4)	Q3 528/2,106 (25.1)	Q4 (High) 530/2,106 (25.2)	5
						Census Tract	% With High School Diplomas	Q1 (Low) 527/2,106 (25.0)	Q2 529/2,106 (25.1)	Q3 524/2,106 (24.9)	Q4 (High) 526/2,106 (25.0)	
Vaz- Luis *et al.* (2014) (269)	USA	SEER-Medicare	28% of US Population 06/2005 - 12/2009 Age ≥ 66 Stage I - III n = 2,028	Trastuzumab n = 2,028 (100.0)	None	Census Tract	Median Household Income	Q1 (Low) 486/2,028 (24.0)	Q2 522/2,028 (25.7)	Q3 508/2,028 (25.1)	Q4 (High) 512/2,028 (25.3)	5
						Census Tract	% With a High School Diploma	Q1 (Low) 503/2,028 (24.8)	Q2: 508/2,028 (25.1)	Q3 508/2,028 (25.1)	Q4 (High) 509/2,028 (25.1)	
Reeder - Hayes *et al.* (2017) (270)	USA	SEER-Medicare	25% of US Population 2005 - 2011 Age ≥ 66 Stage I - III n = 1,077 n (PSM) = 416	Trastuzumab (With Doxorubicin, Cyclophosphamide & Paclitaxel or Docetaxel & Carboplatin Regimens) n = 1,077 (100.0) n (PSM) = 416 (100.0)	None	Census Tract	Residents Living Below the Poverty Line (%)	≥ 20% (Low) 141/1,077 (13.1) PSM 53/416 (12.7)	10% - 19.99% 288/1,077 (26.7) PSM 106/416 (25.5)	5% - 9.99% 288/1,077 (26.7) PSM 117/416 (28.1)	0% - 4.99% (High) 360/1,077 (33.4) PSM 140/416 (33.6)	5

Appendix 2.8 Continued

Sampling Frame						SES		Utilisation by SES Grouping (Number, %)					
Study	**Country**	**Data Source**	**Study Population[a]**	**Predictive Biomarker Test/Biological and Precision Medicine Overall Utilisation (Number, %)**	**Comparator**	**Unit**	**Measure**	**Lowest SES Group**			**Highest SES Group**		**QA**
Chavez-MacGregor *et al.* (2015) (271)	USA	SEER-Medicare & Texas Cancer Registry-Medicare Linked Databases	28% of US Population and Texas State 2005 - 2009 Age ≥ 66 Stage I - III n = 2,203	Trastuzumab n = 2,203 (100.0)	None	NR	Education	4th (Low) 561/2,203 (25.5)	3rd 545/2,203 (24.7)	2nd 547/2,203 (24.8%)	1st (High) 550/2,203 (25.0)		4.5
						NR	Poverty	1st (Low) 550/2,203 (25.0)	2nd 546/2,203 (24.8)	3rd 546/2,203 (24.8)	4th (High) 561/2,203 (25.5)		
Lu *et al.* (2013) (272)	Australia	Medicare Australia (Administering Body of the Nationally Funded Herceptin Program, PBS and MBS)	Women Enrolled in the Herceptin Program 12/2001 – 03/2010 Stage: Metastatic HE2+ Cancer N = 3,418	Trastuzumab n = 3,418 (100.0)	None	Census	Australia's IRSD (Includes Income, Education Attainment & Unemployment Information)	Q1 (Low) 426/3,418 (12.5)	Q2 787/3,418 (23.0)	Q3 644/3,418 (18.8)	Q4 789/3,418 (23.1)	Q5 (High) 772/3,418 (22.6)	3.5
Melanoma: Novel Anti-Cancer Therapies													
Krimphove *et al.* (2019) (273)	USA	NCDB	1,500 CoC Accredited High & Low Immunotherapy Prescribing Hospitals 2011 - 2015 Age ≥ 20 Stage: IV n = 1,863	Immunotherapy n = 1,863 (100.0)	None	Zip Code	Median Household Income	< $37,000 (Low) 215/1,863 (11.5)	$38,000 - $47,999 412/1,863 (22.1)	$48,000 - $62,999 516/1,863 (27.7)	≥ $63,000+ (High) 714/1,863 (38.3)		6.5
						Zip Code	% of Adults Without a High School Diploma	≥ 21% (Low) 258/1,863 (13.8)	13% - 20% 422/1,863 (22.7)	7% - 12% 637/1,863 (34.2)	< 7% (High) 542/1,863 (29.1)		

Appendix 2.8 Continued

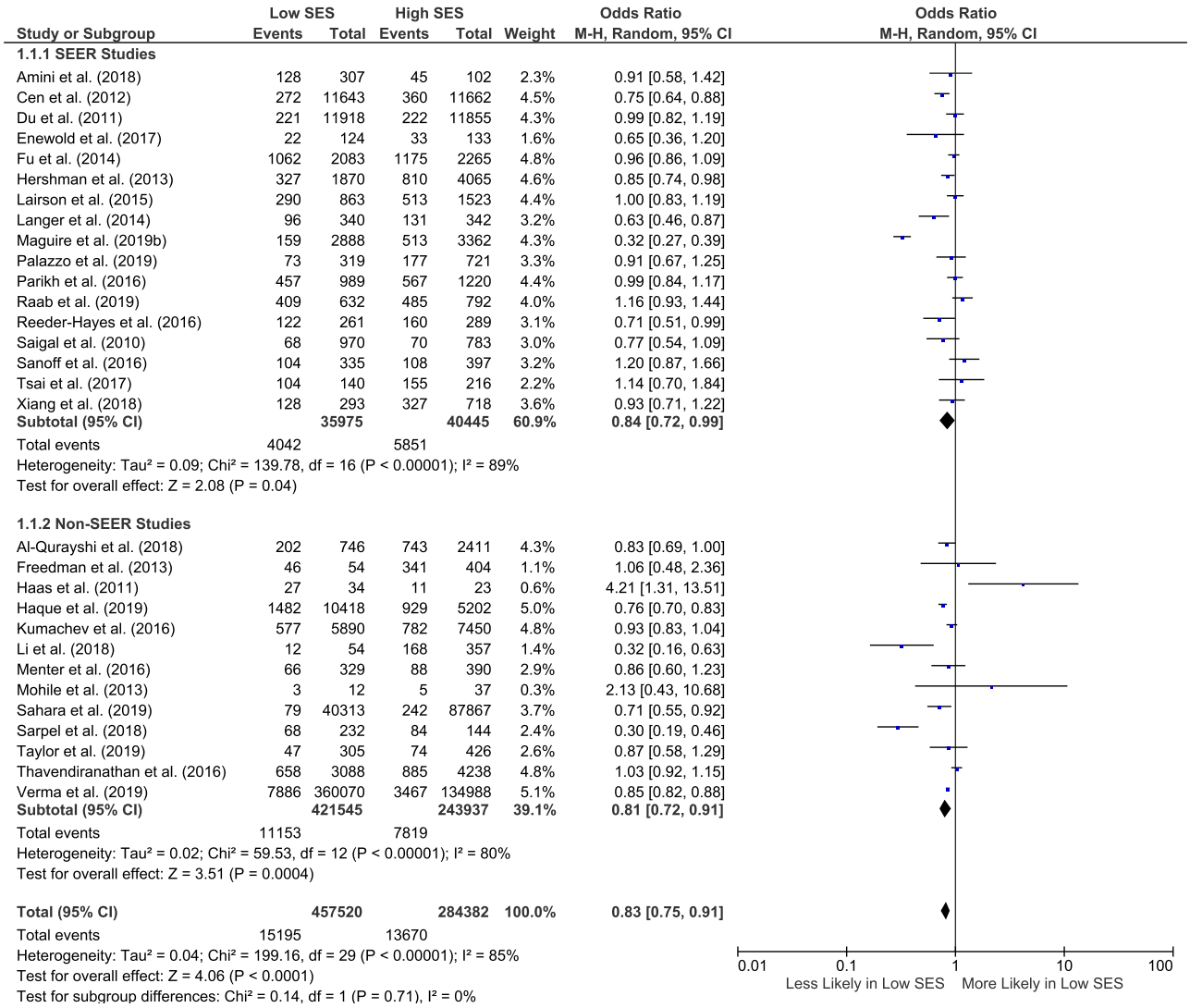
Sampling Frame						SES		Utilisation by SES Grouping (Number, %)				
Study	**Country**	**Data Source**	**Study Population[a]**	**Predictive Biomarker Test/Biological and Precision Medicine Overall Utilisation (Number, %)**	**Comparator**	**Unit**	**Measure**	**Lowest SES Group**			**Highest SES Group**	**QA**
Colon Cancer: Novel Anti-Cancer Therapies												
Zheng *et al.* (2014) (274)	USA	SEER - Medicare	2003 - 2007 Age ≥ 66 Stage: Metastatic n = 7,895	Chemotherapy & Biologics n = 1,260 (16.0)	No Treatment/ Fluorouracil, Capecitabine, Floxuridine, Leucovorin & levoleucovorin (5FU/LV)/ Oxaliplatin, Irinotecan or Both	Zip Code	Household Median Income	Biologics Mean: $51,000 SD $23,000	No Treatment Mean: $48,000 SD $23,000	5FU/LV Mean: $51,000 SD $24,000	Oxaliplatin Irinotecan or Both Mean: $52,000 SD $25,000	**5**
Hepatocellular Cancer: Novel Anti-Cancer Therapies												
Kwan *et al.* (2018) (275)	USA	SEER - Medicare	26% of the US Population 01/2007 - 12/2011 Age ≥ 65 Stage I - IV n = 1,017	Sorafenib n = 369 (36.3)	Embolization	Census Tract	Median Income	Sorafenib Median: $26,607 SD $22,558 P Value: **< 0.001** PSM: Median: $49,445	Embolization Median: $51,774 SD $25,273 PSM: Median: $48,761 Standardized Difference = 3%			7
Renal Cell Carcinoma: Novel Anti-Cancer Therapies												
Li *et al.* (2019) (276)	USA	SEER-Medicare	20 US Registries (28% of US Population) 2000 - 2013 Stage IV n = 1,015	Targeted Therapy (Sorafenib, Sunitinib, Temsirolimus, Everolimus, Bevacizumab, Pazopanib & Axitinib) n = 641 (63.2)	Non-Targeted Therapy	County level	Per Capita Income	Targeted Therapy Mean: $39,900 SD: $10,700 P Value: < **0.001**	Non Targeted Therapy Mean: $31,400 SD: $9,900			5
						County level	Unemployment Rate Per 1,000 Residents	Targeted Therapy Mean: 39.4 SD 16.1 P Value: < **0.001**	Non Targeted Therapy Mean: 27.0 SD 10.5			

Appendix 2.8 Continued

	Sampling Frame					SES		Utilisation by SES Grouping (Number, %)			
Study	**Country**	**Data Source**	**Study Population[a]**	**Predictive Biomarker Test/Biological and Precision Medicine Overall Utilisation (Number, %)**	**Comparator**	**Unit**	**Measure**	**Lowest SES Group**		**Highest SES Group**	**QA**
Head & Neck Cancer: Novel Anti-Cancer Therapies											
Zandberg *et al.* (2018) (277)	USA	SEER - Medicare	17 Registries (28% of US Population) 2005 - 2011 Squamous Cell Age ≥ 66 n = 2,135	Cetuximab with Concurrent Radiation n = 579 (27.1)	Radiation with Concurrent Cytotoxic Chemotherapy/Radiation	Census Tract	Median Income	Cetuximab with Concurrent Radiation Median: $51,000 Range: $8,000 - $232,000	Radiation with Cytotoxic Chemotherapy Median: $49,000 Range: $9,000 - $250,000	Radiation Median: $46,000 Range: $11,000 - $250,000	5.5

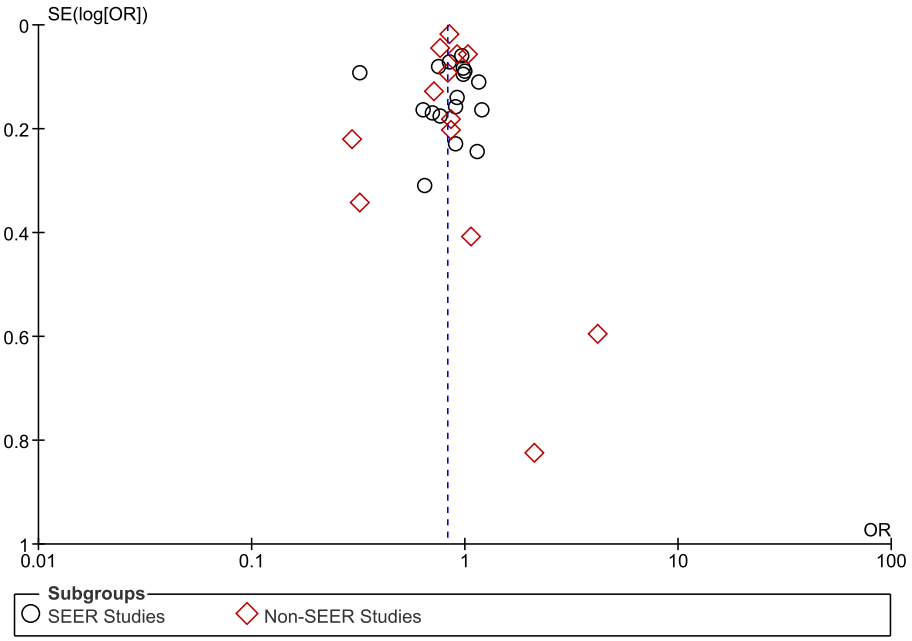
[a]Refers to the total number of patients in the cohort of interest.
Utilisation number (%) reported.
P Values = Significant at $P < 0.05$
Quality appraisal scores range from 0 (lowest) to 10 (highest).
Abbreviations: CoC: Commission on Cancer; IRSD: Index of Relative Socio-economic Disadvantage; MBS: Medicare Benefits Scheme; NCDB: National Cancer Database; NR: Not reported; PBS: Pharmaceutical Benefits Scheme; PSM: Propensity score matching; QA: Quality appraisal; SEER: Surveillance, Epidemiology and End Results Program; SD: Standard deviation; SES: Socio-economic status.

Appendix 2.9 Forest plot of odds of novel anti-cancer therapy utilisation in SEER versus non-SEER registry studies for low compared to high SES (including funnel plot) for Chapter 2



Abbreviations: M-H: Mantel-Haenszel; SEER: Surveillance, Epidemiology, and End Results program; SES: Socio-economic status; 95% CI: 95% Confidence interval.

Appendix 2.9 Continued



Appendix 2.9b Funnel plot of novel anti-cancer therapy utilisation for SEER versus non-SEER studies. OR: Odds ratio; SEER: Surveillance, Epidemiology, and End Results program.

Appendix 2.10 Sensitivity analyses for novel anti-cancer therapy utilisation odds for all cancers (sub-grouped by drug class) for low compared to high SES for Chapter 2

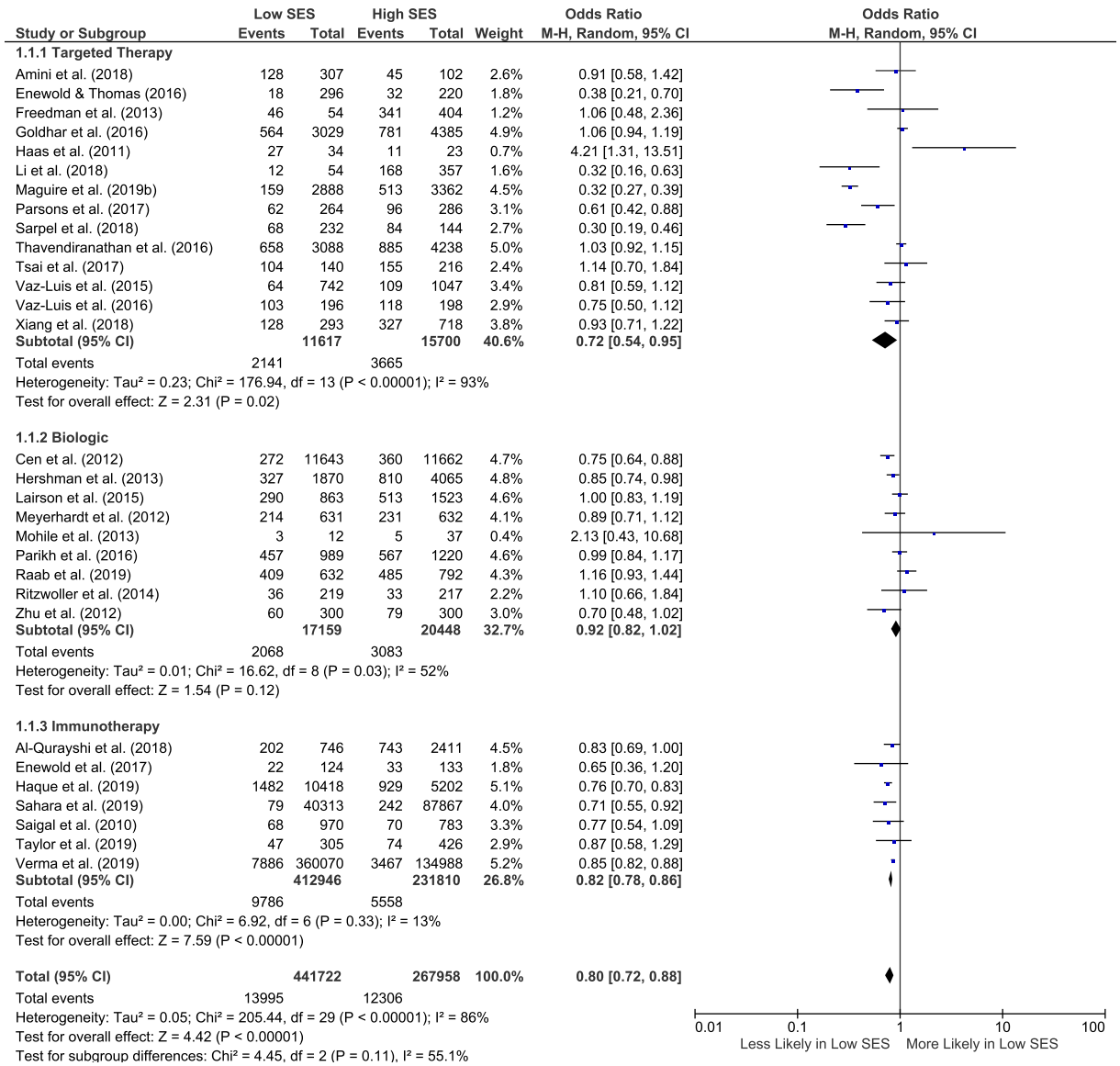


Figure 2.3 (main text) plot but substituting included studies (Du *et al.,* (2011) (286), Kumachev *et al.,* (2016) (294), Palazzo *et al.,* (2019) (325), Reeder-Hayes *et al.,* (2016) (288), Sanoff *et al.,* (2016) (306), Fu *et al.,* (2014) (311), Langer *et al.,* (2014) (299) & Menter *et al.,* (2016) (301)) with those excluded due to overlapping sampling frames.
Abbreviations: M-H: Mantel-Haenszel; SES: Socio-economic status; 95% CI: 95% Confidence interval.

Appendix 2.11 Sensitivity analyses for novel anti-cancer therapy utilisation odds for all cancers (sub-grouped by drug class) for low compared to high SES for Chapter 2

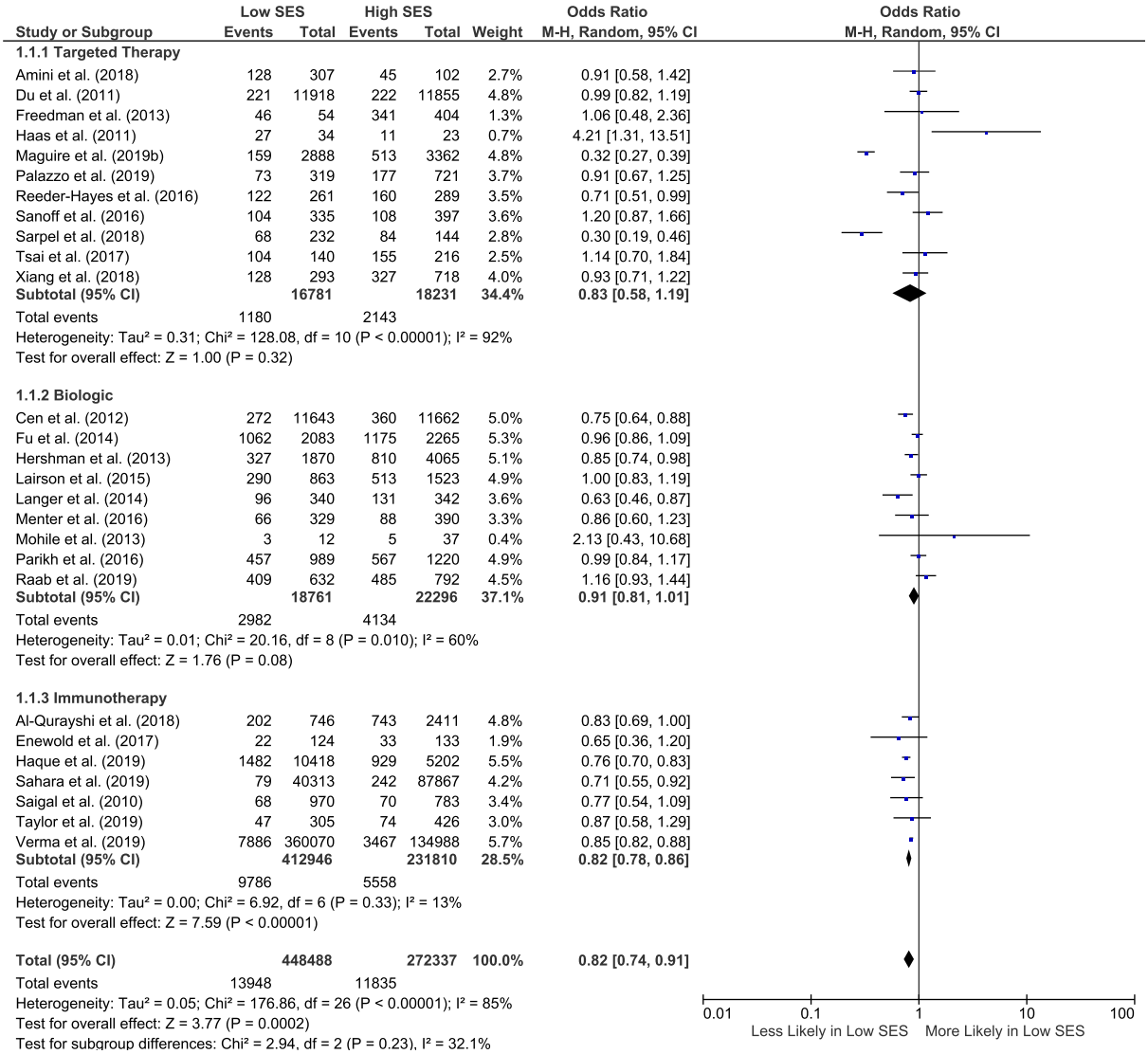
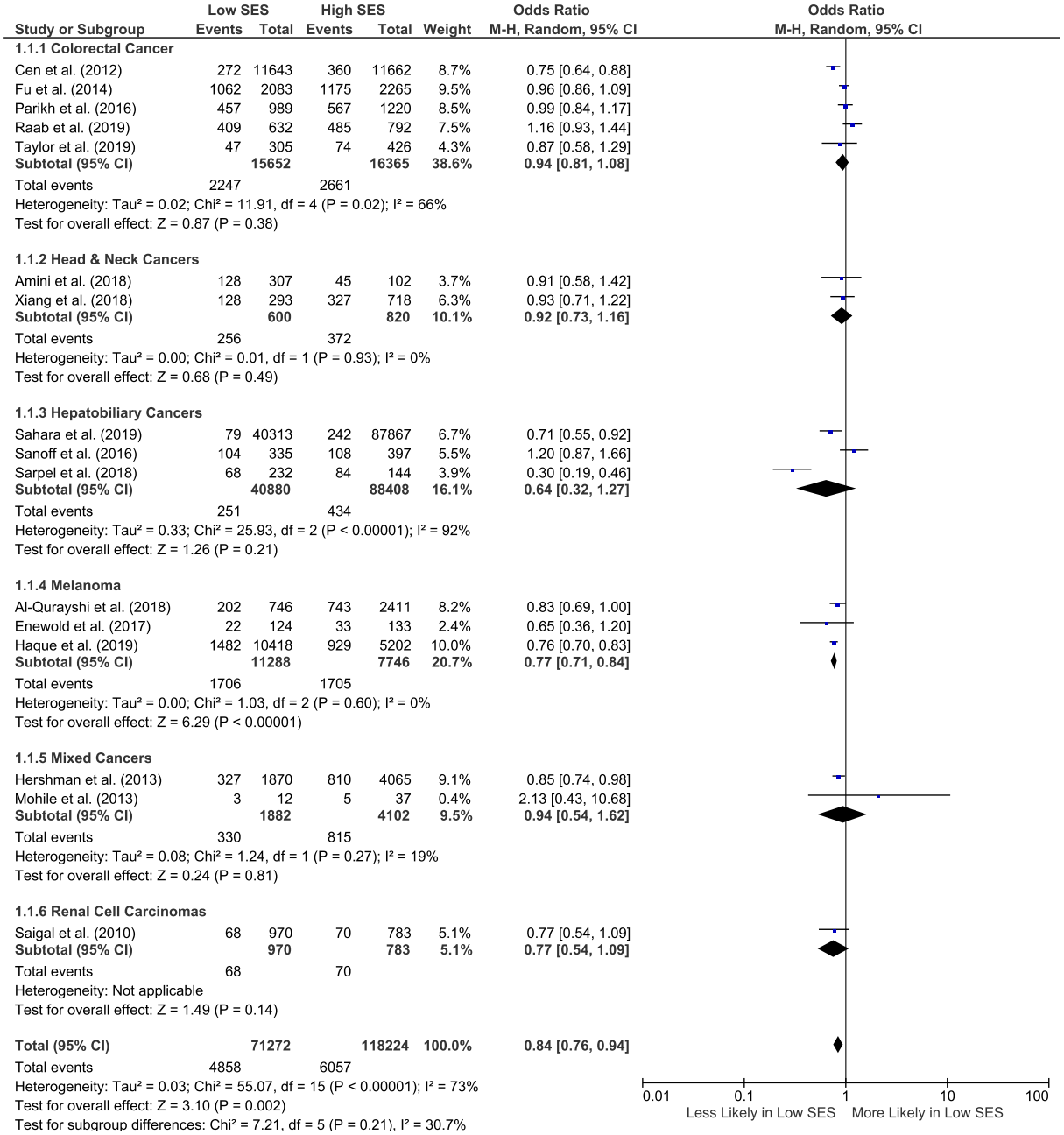


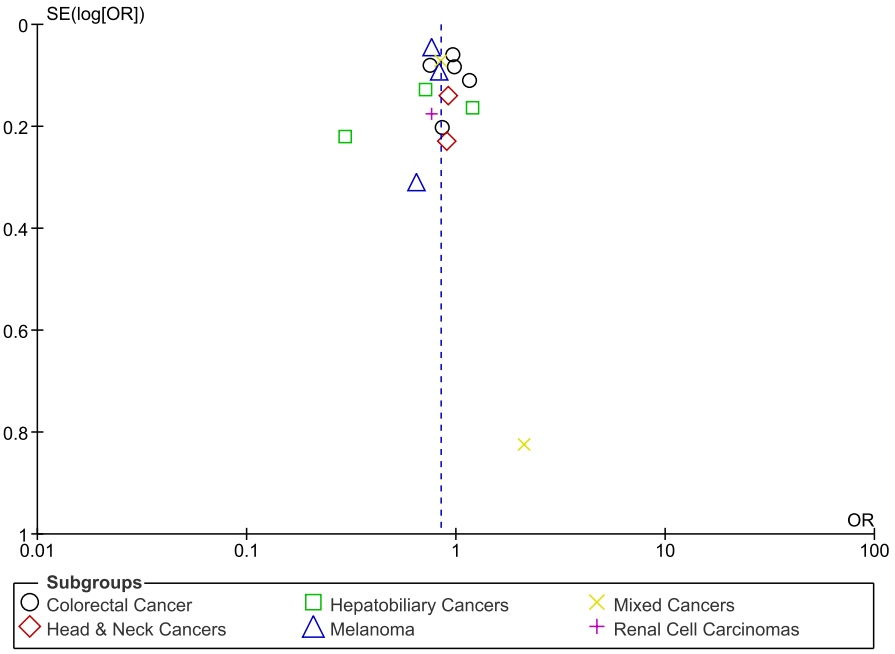
Figure 2.3 plot (main text) but removing the non-USA healthcare setting studies (Kumachev *et al.,* (2016) (294), Li *et al.,* (2018) (296) & Thavendiranathan *et al.,* (2016) (295)).
Abbreviations: M-H: Mantel-Haenszel; SES: Socio-economic status; USA: United States of America; 95% CI: 95% Confidence interval.

Appendix 2.12 Forest plot of odds of novel anti-cancer utilisation (sub-grouped by cancer type) for low compared to high SES (including funnel plot) for all other cancers (apart from breast and lung cancer) for Chapter 2.



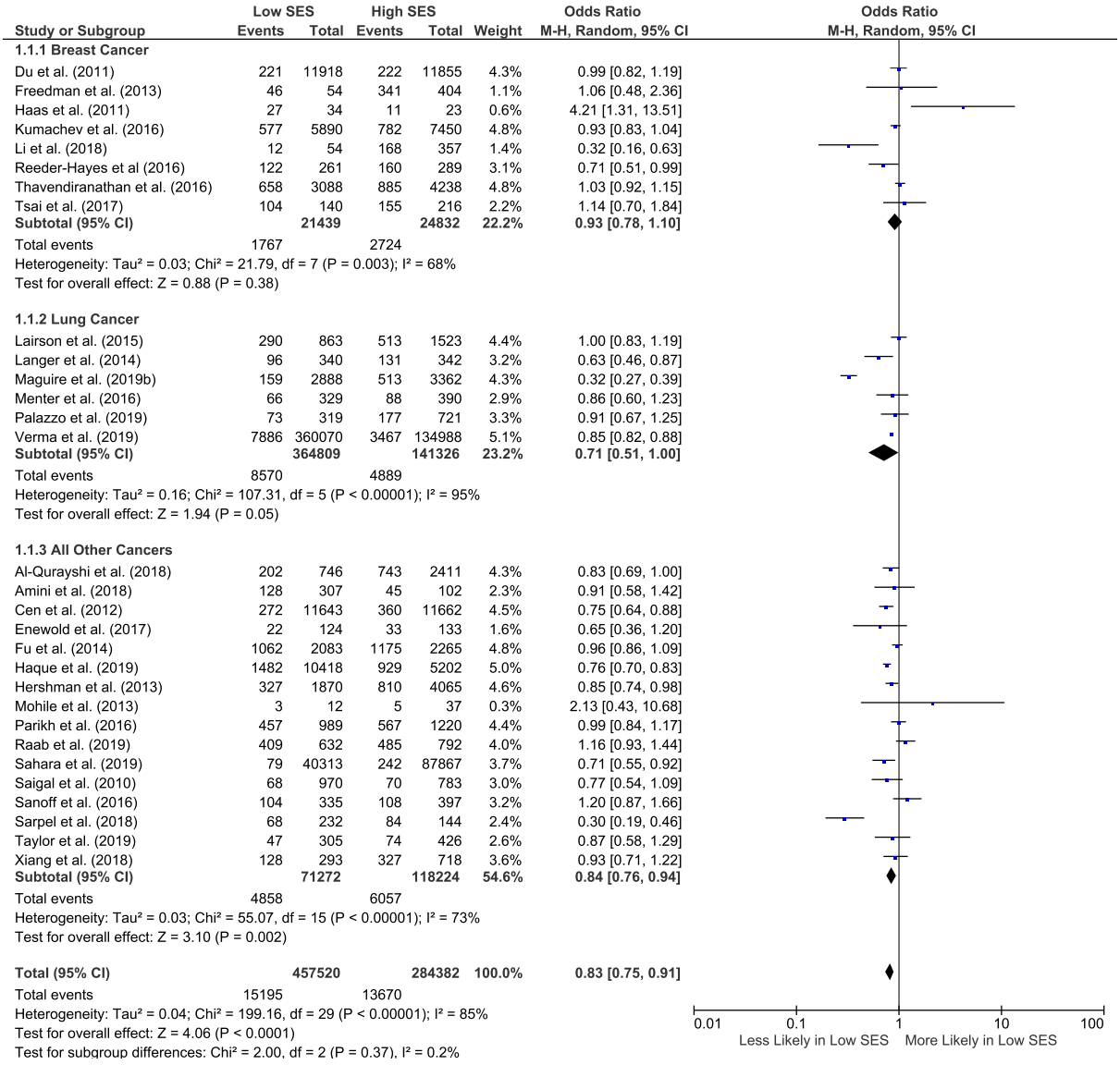
Hershman et al. (2013) (321) refers to breast, colon, and NSCLC cancers.
Mohile et al. (2013) (322) refers to colorectal and NSCLC cancers.
Abbreviations: M-H: Mantel-Haenszel; SES: Socio-economic status; 95% CI: 95% Confidence interval.

Appendix 2.12 Continued



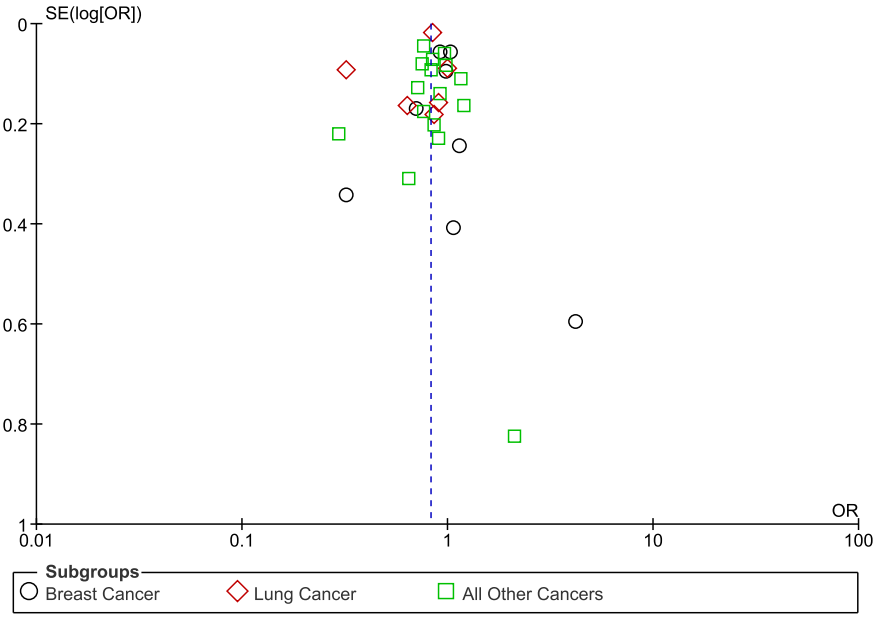
Appendix 2.12b Funnel plot of novel anti-cancer therapy utilisation for all other cancer studies.
Abbreviations: OR: Odds ratio

Appendix 2.13 Forest plot of odds of novel anti-cancer therapy utilisation for all eligible studies (sub-grouped by breast cancer, lung cancer and all other cancers) for low compared to high SES (including funnel plot) for Chapter 2



Abbreviations: M-H: Mantel-Haenszel; SES: Socio-economic status; 95% CI: 95% Confidence interval.

Appendix 2.13 Continued



Appendix 2.13b Funnel plot of novel anti-cancer therapy utilisation for all eligible studies sub-grouped by breast cancer, lung cancer and all other cancers. Abbreviations: OR: Odds ratio.

Appendix 2.14 Funnel plots for all other included meta-analyses with ten plus studies for Chapter 2

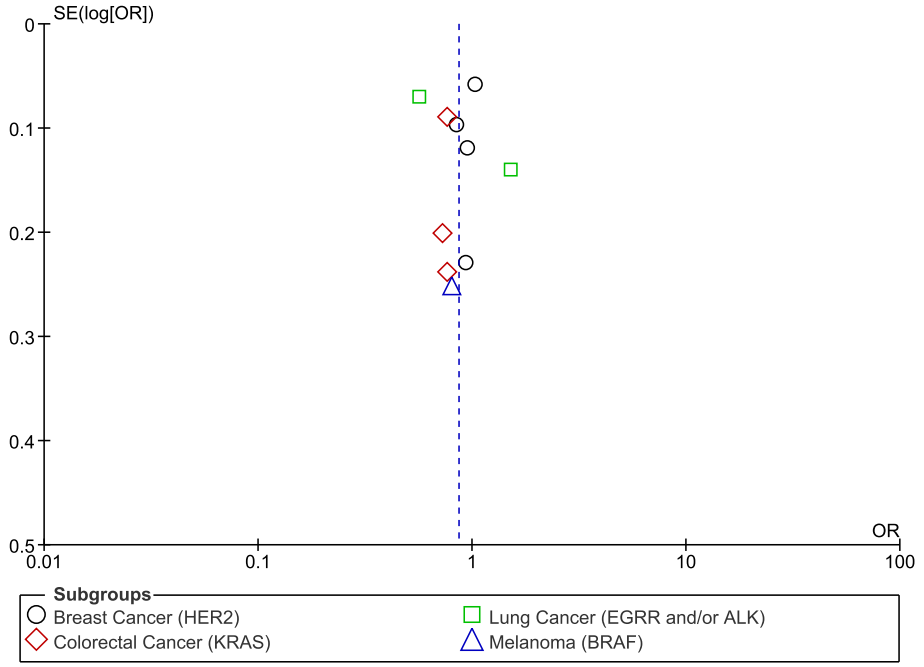


Figure 2.14a: Funnel plot for predictive biomarker test utilisation studies (see Figure 2.2 – main text). Abbreviations: ALK: Anaplastic lymphoma kinase; BRAF: V-raf murine sarcoma viral oncogene homolog B1; EGFR: Epidermal growth factor receptor; HER2: Human epidermal growth factor 2; KRAS: Kirsten rat sarcoma viral oncogene homolog: OR: Odds ratio.

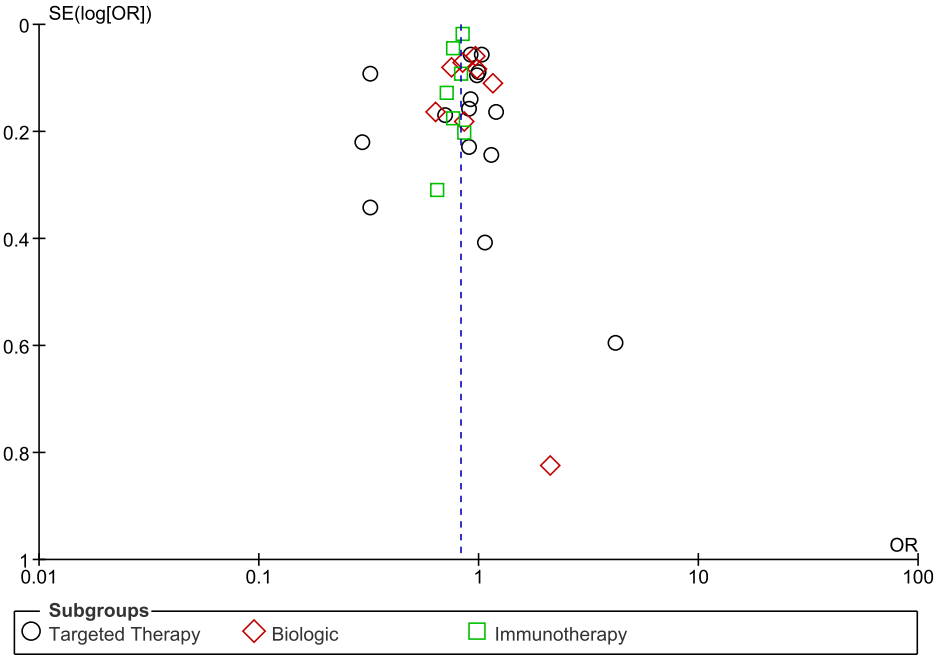


Figure 2.14b: Funnel plot for all cancers' novel anti-cancer therapy utilisation studies (see Figure 2.3 – main text). Abbreviations: OR: Odds ratio.

Appendix 3.1 Overview of studies reporting national SACT dataset as a primary focus of interest for Chapter 3

Study	Dataset(s)	Study Population	Number with a SACT Record	Novel Anti-Cancer Therapies	Findings
Board *et al.* (2021) (403)	NCRD SACT HES	Melanoma 04/2010 – 12/2017 Stage: III/IV (Metastatic) n = 95,259	n = 5,465	First Line Immune Checkpoint Inhibitors n = 2,322 Ipilimumab n =724 Pembrolizumab n = 1,174 Nivolumab n = 52 Ipinivo n =372	Physician first line immune checkpoint inhibitor prescribing is linked to drug access and patient survival/toxicity. RCT outcomes can be achieved in routine care settings with careful patient selection.
Boyle *et al.* (2021) (384)	NBOCA (CRC Database) SACT HES	Colon Cancer 06/2014 – 04/2017 Stage: III n = 10,280	n = 5,109	Adjuvant Chemotherapy n = 1,649	Routine collated national chemotherapy data and administrative hospital data are highly accurate in recording regimen and number of chemotherapy cycles. However, chemotherapy information should be captured in both SACT and HES to avoid under capture.
Boyle *et al.* (2022) (409)	NBOCA SACT HES-APC	Colon Cancer 06/2014 – 04/2017 Stage: III n = 4,147	NR	FOLFOX n = 1,776 CAPOX n = 2,371	Assess the impact of cycle completion rates on survival. Found that completion of all cycles of chemotherapy was associated with improved cancer specific survival.
Boyle *et al.* (2022) (410)	NBOCA HES SACT	Colon Cancer 06/2014 – 12/2017 n = 23,265	NR	Chemotherapy Stage III n = 6,012 Stage IV n = 3,680	Develop and validate a comprehensive coding framework to identify severe acute toxicity in hospital administrative data. The framework captures severe acute toxicity and this can be used to inform clinical decision making.
Elsada *et al.* (2021) (404)	NCRD SACT HES	Multiple Myeloma 01/2013 – 12/2018 n = 28,120	NR	New SACT Line after prior receipt of 3 or more lots including a Proteasome Inhibitor, Immunomodulatory Agent and Anti-CD38 Monoclonal Antibody n = 366	Use real-world data to look at a little studied patient population to highlight poor outcomes. For over 65%, the new SACT line consisted of a pomalidomide-based regimen.

Appendix 3.1 Continued

Study	Dataset(s)	Study Population	Number with a SACT Record	Novel Anti-Cancer Therapies	Findings
Fraser *et al.* (2021) (400)	NCRD SACT	Childhood & Young Adult Cancer 04/2012 – 03/2018 n =145,657	12,272	Single Agent Oral Etoposide N = 115	There are differences in survival between cancers of different histologies - whether these differences are due to differential activity of etoposide or variable natural histories of the cancers concerned is unclear. Most patients died quickly and despite decades of use, there is still no robust data demonstrating the benefit of oral etoposide for survival.
Henson *et al.* (2018) (395)	NCRD RTDS SACT	Lung, Oesophageal, Stomach & Pancreatic Cancer 2013 – 2014 Stage: IV n = 50,232	NR	Chemotherapy n = 24% Chemoradiotherapy n = 11%	There is marked variation in the management of stage IV cancer. Routine data collection could assist surveillance across all cancers to help reduce treatment variation and optimise outcomes in patients with advanced cancer.
Hounsome *et al.* (2022) (407)	NCRD SACT	Diffuse Large B-cell Lymphoma 2013 – 2015 n = 9,186	n = Approx 40% had no SACT recorded or had a treatment outside those of interest	R-CHOP n = 4,392 R-miniCHOP n = 313	Investigation of treatment outcomes in older patients with a focus on the effect of route to diagnosis to outcome (survival). The NHS urgent care pathway is associated with superior survival (equivalent for R-CHOP and R-miniCHOP).
Hovat *et al.* (2022) (412)	COSD SACT ONS	NSCLC 2013 - 2017 Stage: IIIB-IV	NR	PD-L1 ICIs n = 2,305	Consistency of datasets was assessed from England with those from the US was assessed. Found consistent recordings, with any differences relating to timing of ICI approval.
Jones *et al.* (2018) (405)	NLCA SACT HES	SCLC 01/2015 – 12/2015 n = 3,715	NR	Chemotherapy n = 2,235	Administration of chemotherapy was associated with performance status, age, comorbidity, and cancer network. To reduce variation in chemotherapy administration, predictors of 30 day mortality could be used as an adjunct to improve suboptimal patient selection.

Appendix 3.1 Continued

Study	Dataset(s)	Study Population	Number with a SACT Record	Novel Anti-Cancer Therapies	Findings
McDonald *et al.* (2020) (411)	NCRD SACT HES APC CPRD GOLD	Lung cancer 01/2012 – 12/2015 n = 4,070	n = 1,273	Chemotherapy cycles n = 1,273	Of 6,076 chemotherapy cycles, 61% were recorded in SACT on the same day, 8% on a different day and 31% were not recorded in SACT. More than 12.5% of individuals with a HES APC dataset had no SACT chemotherapy record, suggesting that SACT may not capture all chemotherapy cycles administered to patients (up to a third in this instance).
Mintz *et al.* (2019) (406)	SACT HES STAMPEDE Trial	04/2012 – 01/2014 Prostate & Bladder Cancer Single Site (Number 44) n = 8,446	n = 4,156 n = 976 (Randomised after April 2012)	Docetaxel n = 1,573	SACT detected 83/200 STAMPEDE regimens. Quality did not improve between 2012 and 2013. SACT accuracy requires further investigation for trials when a more contemporary STAMPEDE cohort is available. HSPC chemo regimens were not accurately coded in HES or SACT. Although SACT data was seen to enhance the detection of CRPC event, early SACT was not able to identify HSPC event with high accuracy. Some evidence that SACT is a feasible resource to supplement trial collected CRPC therapies but no evidence for accurate HSPC detection due to analysing regimens prior to mandatory collection.
Pathak *et al.* (2017) (392)	SACT Christie Hospital EHR	CRC 04/2011 – 04/2014 Stage: Locally Advanced of Metastatic n = 283	NR	Bevacizumab n = 35% Afibercept n = 23% Cetuximab n = 42% Panitumumab n = NR	Random sample of 20 SACT records. The accuracy of SACT data are dependent on information provided by local hospitals. SACT is potentially a valuable resource for rapidly determining survival outcomes for patients treated with chemotherapy.
Rahman *et al.* (2022) (408)	NOGCA SACT	Oesophagogastric adenocarcinoma 2012 – 2018 n = 4,139	NR	Post Operative Chemotherapy n = 1,593 Platinum-based Triplet Therapy n = 4,004 FLOT Regimens n = 3.3%	Study to investigate the effect of postoperative chemotherapy after surgery. Postoperative chemotherapy in patients treated surgically improved overall survival.

Study	Dataset(s)	Study Population	Number with a SACT Record	Novel Anti-Cancer Therapies	Findings
Wallington *et al.* (2016) (401)	SACT	Breast & Lung Cancer 01/2014 – 12/2014 n = 29,112 (Breast) n = 15,545 (Lung)	NR	SACT therapies for which Cycle Start Date was Present n = 28,364 (Breast), n = 15,045 (Lung)	30 Day mortality increased with age for both patients with breast cancer and NSCLC with curative intent and decreased with age for patients receiving palliative SACT. Hospitals with significantly high 30-day mortality should prompt review.

Abbreviations: APC: Admitted Patient Care; COSD: Cancer Outcomes and Services Dataset; CPRD GOLD: Clinical Practice Research Datalink; CRC: Colorectal Cancer; CRPC: Castrate-resistant prostate cancer; EHR: Electronic Health Record; FLOT: Fluorouracil, leucovorin, oxaliplatin, docetaxel; HES: Hospital Episode Statistics; HES-APC: Hospital Episode Statistics admitted patient care; HSPC: Hormone-sensitive prostate cancer; n: ICIs: Immune checkpoint inhibitors; Number; NBOCA: National Bowel Cancer Audit; NCRD: National Cancer Registry Dataset; NLCA: National Lung Cancer Audit; NOGCA: National Oesophago Gastric Cancer Audit; NR: Not Reported; NSCLC: Non-small Cell Lung Cancer; ONS: Office for National Statistics; PD-L1: Programmed cell death protein ligand 1: RCT: Randomised Controlled Trial; RTDS: Radiotherapy Dataset; R-CHOP: Rituximab, cyclophosphamide doxorubicin, vincristine, and prednisolone; R-miniCHOP: Rituximab and reduced dose CHOP: Systemic Anti-cancer Therapy; SCLC: Small cell lung cancer.

Appendix 3.2 NCRAS Cancer Registration Data Dictionary v3.7
Extracts are shown only for the following v3.7 tables: AV patient, AV tumour, AV treatment, IMD domain and SACT. To view all data dictionary tables, visit: http://ncin.org.uk/collecting_and_using_data/(432). Please note, data variables now available for request may differ to those listed on v3.7 (2019) shown below.

Public Health England

Protecting and improving the nation's health

NCRAS Cancer Registration Data Dictionary v3.7

Access to cancer registration data

The NCRAS data dictionary contains attribute lists that are designed to support the co-production of a data specification for a specific project, ahead of submission to the PHE Office for Data Release (ODR). Prospective applicants are advised to seek expert support from a site, geographic or asset specific analyst to confirm and validate their data requirements. To access the support available through the NCRAS analytical network, please contact ODR@phe.gov.uk

Data availability

The following datasets are available for request for patients who have been diagnosed with ICD 10 C00-97, D00- 48 and registered by NCRAS as part of cancer registration.

Tables	Data is available for:	Temporality of the data (available for follow up)	Available as a discrete dataset (not linked)	Available as linked resource only
Cancer registration (pre 1995) - data dictionary pending, please contract ODR to discuss	Patients diagnosed with from 1 January 1985 - 31 December 1994	01/01/1985 - 31/12/1994	x	
Cancer registration (patient table)	Patients diagnosed with from 1 January 1995 - 31 December 2016	01/01/1995 - 31/12/2016	x	
Cancer registration (tumour table)	Patients diagnosed from 1 January 1995 - 31 December 2016	01/01/1995 - 31/12/2016	x	
Cancer registration (treatment table)	Patients diagnosed from 1 January 1995 - 31 December 2016	01/01/1995 - 31/12/2016	x	
Index of Multiple Deprivation Income domain	Patients diagnosed with from 1 January 1995 - 31 December 2016	01/01/1995 - 31/12/2016		x
Route to diagnosis	Patients diagnosed from 1 January 2006 - 31 December 2016	01/01/2006 - 31/12/2016		x
National Radiotherapy Dataset (RTDS) (Historic)	Patients in receipt of radiotherapy from 1 April 2009 - 31 March 2016	01/04/2009 - 2 months prior to request date		x
Systemic Anti-Cancer Therapy Dataset	Patients diagnosed from 1 April 2012 - 31 March 2018 (where linked data is required, available cohorts will be limited to patients diagnosed up to 31 December 2016)	01/04/2012 - 31/03/2018		x
HES admitted care	Patients diagnosed from 1 January 1995 - 31 December 2016	01/04/2000 - 31/03/2018	*Please contact the NHS Digital Data Access Request Service (DARs) if you require HES data that is not linked to cancer registration*	x
HES outpatient	Patients diagnosed from 1 January 1995 - 31 December 2016	01/04/2004 - 31/03/2018		x
HES accident and emergency	Patients diagnosed from 1 January 1995 - 31 December 2016	01/04/2007 - 31/03/2018		x
Cancer Patient Experience Survey	Wave 1: Patients discharged between 01/01/2010 – 31/03/2010 Wave 2: Patients discharged between 01/09/2011 – 30/11/2011 Wave 3: Patients discharged between 01/09/2012 – 30/09/2012 Wave 4: Patients discharged between 01/09/2013 – 30/11/2013 Wave 5: Patients discharged between 01/04/2015 – 30/06/2015	N/A	x	
Cancer Waiting Times	Patients diagnosed from 1 January 2009 - 31 December 2016	01/01/2009 - 31/12/2017	*Please contact NHS England*	x
Quality of Life of Cancer Survivors in England: (Breast, Colorectal, Prostate, Non-Hodgkin's Lymphoma)	Respondents to a questionnaire distributed to cancer patients in 2011 and sequentially sampled in 2012.	N/A	x	
Lung Cancer Data Audit (LUCADA)	Patients diagnosed between 01/01/2005 - 31/12/2013	01/01/2005 - 31/12/2013	x	
Quality of Life of Colorectal Cancer Survivors in England: Patient Reported Outcome Measures Survey (PROMS)	Respondents to a questionnaire distributed to colorectal cancer patients in January 2013	N/A	x	
Diagnostic Imaging Dataset	Patients diagnosed from 1 January 2013 - 31 December 2015	01/01/2013 - 31/12/2015	*Please contact the NHS Digital Data Access Request Service (DARs) if you require diagnostic imaging data that is not linked to cancer registration data*	x
National Cancer Diagnosis Audit (NCDA)	Subset of patients diagnosed from 1 January 2014 - 31 December 2014 registered at time of diagnosis at one of 439 GP practices in England (~5% of practices) that took part in the NCDA in 2017.	N/A	x	

Appendix 3.2 Continued
Cancer Reg – AV Patient

Introduction
To achieve comprehensive registration for all registerable tumours, the NCRAS brings together data from more than 500 local and regional datasets to build a picture of an individual's treatment from diagnosis. The CAS AV tables includes data on the patient, their diagnosis, tumour and treatment events.

Data item selection
Please be aware that selecting patient identifiable data items – **the data items marked below in red** – will require you to have patient consent, section 251 approval, or the appropriate statutory regulation covering your organisation. Sensitive patient demographic data, event dates and geographies will be treated as de-personalised, unless anonymisation techniques are applied to render the data anonymised to k-3 anonymity. Privacy by design and default principles have also been applied to some of these data to provide a less disclosive alternative. The key explaining how these alternative can be selected is included below:

Key:

Items in red	Personally identifiable data that require specific lawful permissions to access (as outlined above). In requesting access to these variables, please refer to Section 6 of the 'How to complete the ODR Data Request Form (v4) Guidance' for details of evidence that must accompany your application.
	Derived fields that apply privacy by design and can be used for data minimisation

Data item	Field name	Description of field content	Request field (mark required variables with x)	Justification - detail why the field is necessary for your analysis
Pseudonymised patient ID	PATIENTID	Project specific ID for each person	X	For data management.
NHS number	**NHSNUMBER**	**Valid NHS Number or blank.**		
Alias check flag - patient	ALIASFLAG	0,1 (Indicates that this patient record has been deduplicated with another patient and the tumour(s) moved to that other patientid)		
Date of Birth	**BIRTHDATEBEST**	**ddmmyyyy**		
Month of birth	MONTH_DOB	mm		
Year of birth	YEAR_DOB	yyyy		
Date of Birth check flag - patient	BIRTHDATEFLAG	0,1,2,3 (Set to 0 if the date was fully specified, 1 if the month and year of diagnosis are known, but the day was not specified, 2 if the year is fully known, but the month and day are not specified, and 3 if the date was less specific than any of these)		
Sex	SEX	0=Not known, 1=Male, 2=Female, 9=Not specified	X	For descriptive and adjusted analyses.
Ethnicity	ETHNICITY	A = (White) British, B =(White) Irish, C = Any other White background, D = White and Black Caribbean, E = White and Black African, F = White and Asian, G = Any other mixed background, H = Indian, J = Pakistani, K = Bangladeshi, L = Any other Asian background, M = Caribbean , N = African, P = Any other Black background, R = Chinese, S = Any other ethnic group, Z = Not stated, X = Not Known		

Appendix 3.2 Continued

Ethnic group	ETHNICITYNAME	(White) British, (White) Irish, Any other White, background, White and Black Caribbean, White and Black African, White and Asian, Any other mixed background, Indian, Pakistani, Bangladeshi, Any other Asian background, Caribbean , African, Any other Black background, Chinese, Any other ethnic group, Not stated, Not Known	X	For descriptive and adjusted analyses.
Broad ethnic group	Option to group ethnicities (e.g. white/ non-white/ unknown)	Derived as per applicant requirements		
Vital status of the patient	VITALSTATUS	A =Alive, D =Dead, X =Exit posting	X	For survival analyses.
Date of death of the patient	DEATHDATEBEST	ddmmyyyy		
Month of death of the patient	MONTH_DOD	MM		
Year of death of the patient	YEAR_DOD	YYYY		
Days from another event to date to death	Option to provide number of days from another event to death (e.g. days from diagnosis to death)	Derived as per applicant requirements	X	From date of diagnosis to death, or censoring date, depending on vital status (for survival analyses).
Date of death imputed flag	DEATHDATEFLAG	0,1,2,3 (Set to 0 if the date was fully specified, 1 if the month and year of diagnosis are known, but the day was not specified, 2 if the year is fully known, but the month and day are not specified, and 3 if the date was less specific than any of these)		
Embarkation flag	EMBARKATION	Y or blank		
Date of embarkation	EMBARKATIONDATE	ddmmyyyy		
Month of embarkation	Month of embarkation	mm		
Year of embarkation	Year of embarkation	yyyy		
Days from another event to embarkation	Option to provide number of days from another event to embarkation (e.g. days from diagnosis to embarkation	Derived as per applicant requirements		
As provided with death notification	DEATHCAUSECODE_1A	Text – no validation	X	For survival analyses.
As provided with death notification	DEATHCAUSECODE_1B	Text – no validation	X	For survival analyses.
As provided with death notification	DEATHCAUSECODE_1C	Text – no validation	X	For survival analyses.
As provided with death notification	DEATHCAUSECODE_2	Text – no validation	X	For survival analyses.
As provided with death notification	DEATHCAUSECODE_UNDERLYING	Text – no validation	X	For survival analyses.
Code of the location (type) where the patient died, e.g. patients home, hospice etc.	DEATHLOCATIONCODE	1, 2, 3, 4, 5, 6, X, blank		
Description of the location (type) where the patient died, e.g. patients home, hospice etc.	DEATHLOCATIONDESC	CARE HOME, HOSPICE NOS, HOSPITAL, NHS HOSPICE / SPECIALIST PALLIATIVE CARE UNIT, NURSING HOME, OTHER, PRIVATE HOME, UNKNOWN, VOLUNTARY HOSPICE / SPECIALIST PALLIATIVE CARE UNIT, blank		
Code of institution at which death takes place	SITECODEOFDEATH	Valid institution code		
Pseudonymised code of institution at which death takes place	SITECODEOFDEATH (pseudonymised)			
Indicates whether a post-mortem took place	POSTMORTEM	8, 9, N, Y, blank		
Count of every tumour assigned to this PatientID.	TUMOURCOUNT	Number		
Count of every tumour assigned to this PatientID in range C00-97 excl C44	BIGTUMOURCOUNT	Number	X	To determine previous tumour numbers for propensity matching.

********** End of sheet **********

Appendix 3.2 Continued
Cancer Reg – AV Tumour

Introduction
To achieve comprehensive registration for all registerable tumours, the NCRAS brings together data from more than 500 local and regional datasets to build a picture of an individual's treatment from diagnosis. The CAS AV tables includes data on the patient, their diagnosis, tumour and treatment events.

Data item selection
Please be aware that selecting patient identifiable data items – **the data items marked below in red** – will require you to have patient consent, section 251 approval, or the appropriate statutory regulation covering your organisation. Sensitive patient demographic data, event dates and geographies will be treated as de-personalised, unless anonymisation techniques are applied to render the data anonymised to k-3 anonymity. Privacy by design and default principles have also been applied to some of these data to provide a less disclosive alternative. The key explaining how these alternative can be selected is included below:

Key:

Items in red	Personally identifiable data that require specific lawful permissions to access (as outlined above). In requesting access to these variables, please refer to Section 6 of the 'How to complete the ODR Data Request Form (v4) Guidance' for details of evidence that must accompany your application.
	Derived fields that apply privacy by design and can be used for data minimisation

Data item	Field name	Description of field content	Request field (mark required variables with x)	Justification - detail why the field is necessary for your analysis
Pseudonymised tumour ID	TUMOURID	Project specific ID for each tumour	X	For data management.
Pseudonymised patient ID	PATIENTID	Project specific ID for each person	X	For data management.
NHS Number	**NHSNUMBER**	**Valid NHS Number or blank.**		
Date of Birth	**BIRTHDATEBEST**	**ddmmyyyy**		
Month of birth	MONTH_DOB	MM		
Year of birth	YEAR_DOB	YYYY		
Age at diagnosis	AGE	Number or blank		
Age at diagnosis in 5 year age bands (0-4 etc.)	FIVEYEARAGEBAND	0 - 4 YRS | 5 - 9 YRS | 10 - 14 YRS | 15 - 19 YRS | 20 - 24 YRS | 25 - 29 YRS | 30 - 34 YRS | 35 - 39 YRS | 40 - 44 YRS | 45 - 49 YRS | 50 - 54 YRS | 55 - 59 YRS | 60 - 64 YRS | 65 - 69 YRS | 70 – 74 YRS | 75 - 79 YRS | 80 - 84 YRS | Blank)	X	For descriptive and adjusted analyses.
Sex	SEX	0=Not known, 1=Male, 2=Female, 9=Not specified.		
Postcode at Diagnosis	**POSTCODE**	**Postcode-7 format.**		
Outward postcode	POSTCODE_OUTWARD	The area and district component of the Postcode		
Broader geographic area/ IMD quintile	Option to provide geography as deprivation score or aggregate to larger geographic areas such as MSOA or	Derived as per applicant requirements		
Ethnicity	ETHNICITY	A = (White) British, B =(White) Irish, C = Any other White background, D = White and Black Caribbean, E = White and Black African, F = White and Asian, G = Any other mixed background, H = Indian, J = Pakistani, K = Bangladeshi, L = Any other Asian background, M = Caribbean , N = African, P = Any other Black background, R = Chinese, S = Any other ethnic group, Z = Not stated, X = Not Known		
Broad ethnic group	Option to group ethnicities (e.g. white/ non-white/ unknown)	Derived as per applicant requirements		
Earliest date when the diagnosis may have taken place	DIAGNOSISDATE1	ddmmyyyy		
Latest date when the diagnosis may have taken place	DIAGNOSISDATE2	ddmmyyyy		
Diagnosis date	DIAGNOSISDATEBEST	ddmmyyyy	X	For current and previous tumours. Required for analysis of temporal trend and propensity matching.
Month of diagnosis	DIAGNOSISMONTH	mm	X	For analysis of temporal trends (if exact date unavailable).
Year of diagnosis	DIAGNOSISYEAR	yyyy	X	For analysis of temporal trends (if exact date unavailable).
Days from another event to date to diagnosis	Option to provide number of days from another event to diagnosis (e.g. days from birth to diagnosis)	Derived as per applicant requirements		
Date of diagnosis imputed flag	DIAGNOSISDATEFLAG	A flag set to inform if any part of the diagnosis date has been imputed		
Financial year of diagnosis	FINANCIALYEAR	yyyy		
Basis of diagnosis of the tumour	BASISOFDIAGNOSIS	Non-microscopic: 0 = Death certificate 1 = Clinical: Diagnosis made before death without (2-7) 2 = Clinical investigation: Includes all diagnostic techniques without a tissue diagnosis 4 = Specific tumour markers: Includes biochemical and/or immunological markers which are site specific Microscopic: 5 = Cytology: Examination of cells whether from a primary or secondary site, including fluids aspirated using endoscopes or needles. Also including microscopic examination of peripheral blood films and trephine bone marrow aspirates6 = Histology of a metastases: Includes autopsy specimens 7 = Histology of a primary tumour: Includes all cutting and bone marrow biopsies. Also includes autopsy specimens of a primary tumour 9 = Unknown, e.g. PAS or HISS record only		
Diagnosis death certificate only	DCO	Y = Yes, N = No	X	For survival analyses.

Appendix 3.2 Continued

Site of neoplasm (4-character ICD-10-O2 code)	SITE_ICD10_O2	Valid 4 digit ICD-10 codes in the range C00-D48 plus D76, E85, O01, Q85 or blank	X	For present tumour and previous invasive cancers (bar non-melanoma skin cancer). Required for adjusted analyses to differentiate breast and lung cancers, as well as to estbalish if the patient has a tumou
Site of neoplasm (3-character ICD-10-O2 code)	SITE_ICD10_O2_3CHAR	Valid 3 digit ICD-10 codes in the range C00-D48 plus D76, E85, O01, Q85 or blank		
Site of the cancer	SITE_CODED	Site of the cancer, in the coding system that the tumour was originally coded in.		
Description of the code in SITE_CODED	SITE_CODED_DESC	Text description of the code in SITE_CODED		
3 digit version of SITE_CODED	SITE_CODED_3CHAR	Three digit version of site_coded		
The coding system used to register the tumour	CODING_SYSTEM	1 = ICD-8, 2 = ICD-9, 3 = ICD-10/O-2, 4 = ICD-10/O-3, 5 = ICD-O-3, 6 = ICD-7, 7 = ICD-8pre1971, 8 = ICD-O-2, 9 = ICD-O, 10 = ICD-O-3 (2011), 11 = ICD-10rev4/O-2, 12 = MOTNAC, 14 = SNOMED/OTCR), 15 = SNOMED/O-1, 16 = SNOMED/O-2, 17 = SNOMED/O-3		
Description of coding system used in registration	CODING_SYSTEM_DESC	TBC		
Morphology	MORPH_CODED	TBC		
Morphology of the cancer, in the ICD-10-O2 system	MORPH_ICD10_O2	Number 8000-9990 or blank	X	For adjusted analyses.
Behaviour of the cancer, in the ICD-10-O2 system	BEHAVIOUR_ICD10_O2	0, 1,2,3,5,6,9,XXX,XXXX, blank		
Numeric behaviour code	BEHAVIOUR_CODED	0 = Benign, 1 = In situ, 2 = Malignant, 3 =Malignant, metastatic / secondary site, 5 = Malignant, uncertain whether primary or metastatic, 6 = Micro-invasive, 9 = Uncertain		
Description of behaviour code	BEHAVIOUR_CODED_DESC	Description of behaviour code		
Histology code	HISTOLOGY_CODED	Histology code		
Description of histology code	HISTOLOGY_CODED_DESC	Text – no validation		
Grade of tumour	GRADE	GX = Grade of differentiation is not appropriate or cannot be assessed G1 = Well differentiated G2 = Moderately differentiated G3 = Poorly differentiated G4 = Undifferentiated / anaplastic	X	For adjusted analyses.
Size of the largest dimension of the tumour, in mm	TUMOURSIZE	Number or blank		
Number of nodes excised	Nodes_excised_new	Number or blank		
Number of nodes involved	nodes_involved_new	Number or blank		
Laterality	LATERALITY	L = Left, R = Right, M = Midline, B = Bilateral, 8 = Not applicable, 9 = Not Known		
Multifocal	MULTIFOCAL	N= No, Y = Yes, 8 = Not applicable, 9 = Not known		
Oestrogen receptor status of the tumour	ER_STATUS	N = negative, P = positive, X = not performed	X	For descriptive and adjusted analyses.
Oestrogen receptor score of the tumour.	ER_SCORE	ER Allred score (range 0, 2-8)	X	For descriptive and adjusted analyses.
Progesterone receptor status of the tumour	PR_STATUS	N = negative, P = positive, X = not performed	X	For descriptive and adjusted analyses.
Progesterone receptor score of the tumour	PR_SCORE	ER Allred score (range 0, 2-8)	X	For descriptive and adjusted analyses.
HER2 status of the tumour	HER2_STATUS	N = negative, P = positive, X = not performed	X	For descriptive and adjusted analyses.
Nottingham Prognostic Index Score	NPI	Number (two decimal places) or blank		
Dukes' stage	DUKES	A = Dukes' A: Tumour confined to wall of bowel, nodes negative B = Dukes' B: Tumour penetrates through the muscularis propria to involve extramural tissues, nodes negative C1 = Dukes' C1: Metastases confined to regional lymph nodes (node/s positive but apical node negative) C2 = Dukes' C2: Metastases present in nodes at mesenteric artery ligature (apical node positive) D = Dukes D: Metastatic spread outside the operative field 99 = Not Known		
FIGO stage	FIGO	0, 1, 1a, 1a1, 1a2, 1b, 1b1, 1b2, 1c, 1c1, 1c2, 1c3, 2, 2a, 2a1, 2a2, 2b, 2c, 3, 3a, 3b, 3c, 3c1, 3c2, 4, 4a, 4b, I, IA, IA1, IA2, IB, IB1, IB2, IC, II, IIA, IIA2, IIB, IIC, III, IIIA, IIIB, IIIC, IIIC1, IIIC2, IV, IVA, IVB, blank		
Clark's stage	CLARKS	1, 2, 3, 4, 5, blank		
Breslow thickness of tumour	BRESLOW	Number or range, x, or blank		
Gleason primary pattern	GLEASON_PRIMARY	1-5, 8 = not applicable		
Gleason secondary pattern	GLEASON_SECONDARY	1-5, 8 = not applicable		
Gleason tertiary pattern	GLEASON_TERTIARY	1-5, 8 = not applicable		
Combined Gleason primary and secondary scores	GLEASON_COMBINED	2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, blank		
T stage (pre-treatment)	T_IMG	UICC code	X	For descriptive and adjusted analyses (to account for confounding).
N stage (pre-treatment)	N_IMG	UICC code	X	For descriptive and adjusted analyses (to account for confounding).
M stage (pre-treatment)	M_IMG	0 = no distant metastasis 1, 1a, 1b, 1c, 1e = distant metastasis X = unknown	X	For descriptive and adjusted analyses (to account for confounding).
Stage at diagnosis derived from imaging	STAGE_IMG	Text		
System used to record imaging stage at diagnosis	STAGE_IMG_SYSTEM	5 = 5th, 6 = 6th, 7 = 7th, 20 = UICC 5, 21 = UICC 6, 22 = UICC 7, 23 = AJCC 7, 24 =Unknown		
T stage (pathology)	T_PATH	UICC code		
N stage (pathology)	N_PATH	UICC code		
M stage (pathology)	M_PATH	0, 1, 1a, 1b, 1c, 1e, 2, 3, 4, 9, X, blank		
Pathological stage at diagnosis	STAGE_PATH	0, 0A, 0IS, 1, 1A, 1A1, 1A2, 1B, 1B1, 1B2, 1C, 1E, 2, 2A, 2B, 2C, 2E, 3, 3A, 3B, 3C, 3E, 4, 4A, 4B, 4C, 5, 6, ?, U, X, blank		
System used to record pathological stage at diagnosis	STAGE_PATH_SYSTEM	5, 6, 7, 20, 21, 22, 23,24, blank		

Appendix 3.2 Continued

Pathological stage at diagnosis (pre-treatment)	STAGE_PATH_PRETREATED	Y = Yes, N = No		
T stage flagged by the registry as the 'best' T stage	T_BEST	UICC code	X	For descriptive and adjusted analyses.
N stage flagged by the registry as the 'best' N stage	N_BEST	UICC code	X	For descriptive and adjusted analyses.
M stage flagged by the registry as the 'best' M stage	M_BEST	UICC code	X	For descriptive and adjusted analyses.
Best 'registry' stage at diagnosis of the tumour	STAGE_BEST	0, 0A, 0IS =Stage 0 1, 1A, 1A1, 1A2, 1B, 1B1, 1B2, 1C, 1E = Stage 1 2, 2A, 2A1, 2A2, 2B, 2C, 2E, 2S = Stage 2 3, 3A, 3B, 3C, 3E, 3S = Stage 3 4, 4A, 4B, 4C, 4S = Stage 4 6 = not stageable ? = insufficient information U = unstageable, X = not staged	X	For descriptive and adjusted analyses.
System used to record best registry stage at diagnosis	STAGE_BEST_SYSTEM	5 = 5th, 6 = 6th, 7 = 7th, 20 = UICC 5, 21 = UICC 6, 22 = UICC 7, 23 = AJCC 7, 24 =Unknown		
Code for the place where the diagnosis episode took place	DIAGNOSISPROVIDER_CODE	Valid provider code		
Pseudonymised diagnosis provider code	DIAGNOSISPROVIDER_CODE (pseudonym	To be derived on request		
Description of DIAGNOSISPROVIDER_CODE	DIAGNOSISPROVIDER_NAME	Text - no validation		
Code for the Trust at diagnosis	DIAGNOSISTRUST_CODE	Valid trust code		
Pseudonymised diagnosis trust code	DIAGNOSISTRUST_CODE (pseudonymised	To be derived on request	X	For descriptive and adjusted analyses. Pseudonymised data is sufficient.
Name of the trust at diagnosis	DIAGNOSISTRUST_NAME	Text - no validation		
Tumour registration status	STATUSOFREGISTRATION	F= registration is final; P= provisional		
Excision margin	EXCISIONMARGIN	01 = Excision margins are clear (distance from margin not stated) 02 = Excision margins are clear (tumour >5mm from the margin) 03 = Excision margins are clear (tumour >1mm but less than or equal to 5mm from the margin 04 = Tumour is less than or equal to 1mm from excision margin, but does not reach margin 05 = Tumour reaches excision margin 06 = Uncertain 07 = Margin not involved =>1mm 08 = Margin not involved <1mm 09 = Margin not involved 1-5mm 98 = Not applicable 99 = Not Known		
Screen detected cancer	SCREENDETECTED	N = No, Y = Yes, 8 = Not applicable, 9 = Not known		
Screening status of the tumour	SCREENINGSTATUSCOSD_CODE	TBC		
Description of SCREENINGSTATUSCOSD_CODE	SCREENINGSTATUSCOSD_NAME	Text - no validation		
Full detailed screening status of the tumour	SCREENINGSTATUSFULL_CODE	TBC		
Description of SCREENINGSTATUSFULL_CODE	SCREENINGSTATUSFULL_NAME	Text - no validation		
Date of first recorded event in treatment table	DATE_FIRST_EVENT	ddmmyyyy		
Month of first recorded event in treatment table	Month of first recorded event in treatment ta	mm		
Year of first recorded event in treatment table	Year of first recorded event in treatment tabl	yyyy		
Days from another event to first recorded event	Option to provide number of days from another event to the first recorded event in the treatment table (e.g. days from diagnosis to first treatment event)	Derived as per applicant requirements		
Trust code of first recorded event in treatment table	TRUSTCODE_FIRST_EVENT	Valid trust code	X	For descriptive and adjusted analyses. Pseudonymised data is sufficient. Not including biopsy.
Pseudonymised trust code of first event	TRUSTCODE_FIRST_EVENT (Pseudonymi	Derived as per applicant requirements		
Name of trust for first recorded event in treatment table	TRUSTNAME_FIRST_EVENT	Text - no validation	X	For descriptive and adjusted analyses. Pseudonymised data is sufficient.
Date of first recorded surgery in treatment table	DATE_FIRST_SURGERY	ddmmyyyy		
Month of first recorded surgery in treatment table	Month of first recorded surgery in treatment	mm		
Year of first recorded surgery in treatment table	Year of first recorded surgery in treatment ta	yyyy		
Days from another event to first recorded surgery in treatment table	Option to provide number of days from another event to the first recorded surgery (e.g. days from diagnosis to first recorded surgery)	Derived as per applicant requirements		
Trust code of first recorded surgery in treatment table	TRUSTCODE_FIRST_SURGERY	Valid trust code		
Pseudonymised trust code of first recorded surgery	TRUSTCODE_FIRST_SURGERY (pseudonymised)	Derived as per applicant requirements		
Name of trust for first recorded surgery in treatment table	TRUSTNAME_FIRST_SURGERY	Text - no validation		
2011 Lower Super Output Area	LSOA11_CODE	ONS code format: X00000000, blank		
2001 Lower Super Output Area	LSOA01_CODE	ONS code format: X00000000, blank		
2011 Middle Super Output Area	MSOA11_CODE	ONS code format: X00000000, blank		
2001 Middle Super Output Area	MSOA01_CODE	ONS code format: X00000000, blank		
Clinical Commissioning Group code (at diagnosis)	CCG_CODE	Code format: 00X, blank	X	For descriptive and adjusted analyses. Pseudonymised data is sufficient.
Name of the Clinical Commissioning Group	CCG_NAME	Text - no validation		
Primary Care Trust code the patient was resident in when the tumour was diagnosed	PCT_CODE	3 digit PCT code, blank		
Name of the Primary Care Trust the patient was resident in when the tumour was diagnosed	PCT_NAME	Text - no validation	X	For descriptive and adjusted analyses. Pseudonymised data is sufficient.
Local Authority Unitary Authority code the patient was resident in when the tumour was diagnosed	LAUA_CODE	00XX UA code	X	For descriptive and adjusted analyses. Pseudonymised data is sufficient.
Name of the Local Authority Unitary Authority the patient was resident in when the tumour was diagnosed	LAUA_NAME	Text - no validation		
Upper tier Local Authority code the patient was resident in when the tumour was diagnosed	UTLA_CODE	00XX UA code, or number, or blank		
Name of the upper tier Local Authority the patient was resident in when the tumour was diagnosed	UTLA_NAME	Text – no validation		

Appendix 3.2 Continued

Upper tier Local Authority code the patient was resident in when the tumour was diagnosed	UTLA_CODE	00XX UA code, or number, or blank		
Name of the upper tier Local Authority the patient was resident in when the tumour was diagnosed	UTLA_NAME	Text – no validation		
Strategic Clinical Network code the patient was resident in when the tumour was diagnosed	SCN_CODE	N44, N50, N51, N52, N53, N54, N55, N56, N57, N58, N59, N60, N61, N95, N96, Z99, blank		
Name of the Strategic Clinical Network the patient was resident in when the tumour was diagnosed	SCN_NAME	Text – no validation		
Cancer network code the patient was resident in when the tumour was diagnosed	CNET_CODE	N01, N02, N03, N06, N07, N08, N11, N12, N20, N21, N22, N23, N24, N25, N26, N27, N28, N29, N30, N31, N32, N33, N34, N35, N36, N37, N38, N39, N95, N96, Z99, blank		
Name of the cancer network the patient was resident in when the tumour was diagnosed	CNET_NAME	Text – no validation		
County code the patient was resident in when the tumour was diagnosed	COUNTY_CODE	11, 12, 16, 17, 18, 19, 21, 22, 23, 24, 26, 29, 30, 31, 32, 33, 34, 36, 37, 38, 40, 41, 42, 43, 44, 45, 47, blank		
Name of the county the patient was resident in when the tumour was diagnosed	COUNTY_NAME	Text – no validation		
Government office region code the patient was resident in when the tumour was diagnosed	GOR_CODE	A, B, D, E, F, G, H, J, K, blank	X	For descriptive and adjusted analyses.
Name of the government office region the patient was resident in when the tumour was diagnosed	GOR_NAME	East Midlands, East of England, London, North East, North West, South East, South West, West Midlands, Yorkshire and The Humber	X	For descriptive and adjusted analyses.
Cancer registry catchment area code the patient was resident in when the tumour was diagnosed	CREG_CODE	Y0801=Thames Cancer Registry, Y0201=Northern & Yorkshire Cancer Registry & Information Service, Y0301=Trent Cancer Registry, Y1201=West Midlands Cancer Intelligence Unit, Y0401=Eastern Cancer Registration & Information Centre, Y1701=North West Cancer Intelligence Service, Y1001=South West Cancer Intelligence Service, Y1101=Welsh Cancer Intelligence & Surveillance Unit, Y0901=Oxford Cancer Intelligence Unit Z9999=blank		
Name of the cancer registry catchment area the patient was resident in when the tumour was diagnosed	CREG_NAME	Eastern Cancer Registration & Information Centre, North West Cancer Intelligence Service, Northern & Yorkshire Cancer Registry & Information Service, Oxford Cancer Intelligence Unit, South West Cancer Intelligence Service, Thames Cancer Registry, Trent Cancer Registry, Welsh Cancer Intelligence & Surveillance Unit, West Midlands Cancer Intelligence Unit		
Country code the patient was resident in when the tumour was diagnosed	CTRY_CODE	11, 12, 16, 17, 18, 19, 21, 22, 23, 24, 26, 29, 30, 31, 32, 33, 34, 36, 37, 38, 40, 41, 42, 43, 44, 45, 47, blank		
Name of the country the patient was resident in when the tumour was diagnosed	CTRY_NAME	Text - no validation		
Cancer registry code which finalised the case and was responsible for sending it to ONS if it was an in-region case	CENTRE	0101, 0201, 0202, 0301, 0302, 0401, 0402, 0403, 0404, 0500, 0600, 0801, 0802, 0901, 1001, 1002, 1201, 1301, 1401, 1501, 1702, NBTR, blank,		
Name of the cancer registry which finalised the case and was responsible for sending it to ONS if it was an in-region case	CENTRENAME	ECRIC BEDFORD, ECRIC CAMBRIDGE, ECRIC IPSWICH, ECRIC NORWICH, FHSA, MERSEY MERSEYSIDE AND CHESHIRE CANCER REGISTRY,		

********** **End of sheet** **********

Appendix 3.2 Continued
Cancer Reg – AV Treatment

Introduction
To achieve comprehensive registration for all registerable tumours, the NCRAS brings together data from more than 500 local and regional datasets to build a picture of an individual's treatment from diagnosis. The CAS AV tables includes data on the patient, their diagnosis, tumour and treatment events.

Data item selection
Please be aware that selecting patient identifiable data items – **the data items marked below in red** – will require you to have patient consent, section 251 approval, or the appropriate statutory regulation covering your organisation. Sensitive patient demographic data, event dates and geographies will be treated as de-personalised, unless anonymisation techniques are applied to render the data anonymised to k-3 anonymity. Privacy by design and default principles have also been applied to some of these data to provide a less disclosive alternative. The key explaining how these alternative can be selected is included below:

Key:

Items in red	Personally identifiable data that require specific lawful permissions to access (as outlined above). In requesting access to these variables, please refer to Section 6 of the 'How to complete the ODR Data Request Form (v4) Guidance' for details of evidence that must accompany your application.
	Derived fields that apply privacy by design and can be used for data minimisation

Data item	Field name	Description of field content	Request field (mark required variables with x)	Justification - detail why the field is necessary for your analysis
Pseudonymised event ID	EVENTID	Project specific ID for each event		
Pseudonymised tumour ID	TUMOURID	Project specific ID for each tumour	X	For data management.
Pseudonymised patient ID	PATIENTID	Project specific ID for each person	X	For data management.
Age at diagnosis	AGE	Number or blank		
Age at diagnosis in 5 year age bands (0-4 etc.)	FIVEYEARAGEBAND	0 - 4 YRS | 5 - 9 YRS | 10 - 14 YRS | 15 - 19 YRS | 20 - 24 YRS | 25 - 29 YRS | 30 - 34 YRS | 35 - 39 YRS | 40 - 44 YRS | 45 - 49 YRS | 50 - 54 YRS | 55 - 59 YRS | 60 - 64 YRS | 65 - 69 YRS | 70 – 74 YRS | 75 - 79 YRS | 80 - 84 YRS | Blank		
Age at diagnosis in x year age bands	Option to provide age in broad categories (e.g. =<45, 46-55, 56-65, >65)	Derived as per applicant requirements		
Sex	SEX	0=Not known, 1=Male, 2=Female, 9=Not specified		
Diagnosis date	DIAGNOSISDATEBEST	ddmmyyyy		
Month of diagnosis	DIAGNOSISMONTH	mm		
Year of diagnosis	DIAGNOSISYEAR	yyyy		
Days from another event to date to diagnosis	Option to provide number of days from another event to diagnosis (e.g. days from birth to diagnosis)	Derived as per applicant requirements		
Number of tumours affected by this event	NUMBER_OF_TUMOURS	Number		
Type of event code	EVENTCODE	01a = Surgery – curative, 01b = Surgery - not curative, 01z = Surgery etc. - type unknown, 02 = Cytotoxic Chemotherapy, 03 = Hormone Therapy, 05 = RT – Teletherapy, 06 =RT – Brachytherapy, 15 = Immunotherapy, 97 = Other Treatment, 99 = Treatment unknown, CTX = CT – Other, IM = Imaging, RTX = RT - Other/NK		
Description of the event	EVENTDESC	Text – no validation		
Date the event took place	EVENTDATE	ddmmyyyy		
Month the event took place	Month of the year the event took place	MM		
Year the event took place	EVENTYEAR	YYYY		
Days from another event to this event	Option to provide number of days from another recorded event to this event (e.g. days from diagnosis to event)	Derived as per applicant requirements		
Treatment provider (organisation code)	PROVIDERCODE	Valid institution code		
Pseudonymised treatment provider code	PROVIDERCODE (pseudonymised)	Derived as per applicant requirements		
Name of the organisation where the event took place	PROVIDERDESC	Text – no validation		

Appendix 3.2 Continued

Code of the NHS Trust where the event took place	TRUST_CODE	Valid Trust code		
Pseudonymised NHS Trust code where the event took place	TRUST_CODE (pseudonymised)	Derived as per applicant requirements		
Name of the NHS Trust where the event took place	TRUST_NAME	Text – no validation		
Consultant code	PRACTITIONERCODE	Valid consultant or GP code		
Consultant code (pseudonymised by default)	PRACTITIONERCODE (pseudonymised)	To be derived for the applicant		
Consultant name	PRACTITIONERDESC	Text – no validation		
Cancer registry catchment area code the patient was resident in when the tumour was diagnosed	CREG_CODE	Y0801=Thames Cancer Registry, Y0201=Northern & Yorkshire Cancer Registry & Information Service, Y0301=Trent Cancer Registry, Y1201=West Midlands Cancer Intelligence Unit, Y0401=Eastern Cancer Registration & Information Centre, Y1701=North West Cancer Intelligence Service, Y1001=South West Cancer Intelligence Service, Y1101=Welsh Cancer Intelligence & Surveillance Unit, Y0901=Oxford Cancer Intelligence Unit Z9999=blank		
Treatment within 6 months of diagnosis - check flag	WITHIN_SIX_MONTHS_FLAG	0 = No, 1 = Yes	X	Receipt of chemotherapy and/or radiotherapy within 6 months - yes/no categorisation for each treatment variable (to establish targeted therapy place in treatment).
Treatment six months from date of diagnosis - check flag	SIX_MONTHS_AFTER_FLAG	0 = No, 1 = Yes		
Operations, procedures and interventions (OPCS-4)	OPCS4_CODE	Valid OPCS4 code		
Name of the operations, procedures and interventions	OPCS4_NAME	Text - no validation		
Radiotherapy code	RADIOCODE	1 = 1 + 2, 2 = 1 + 4, 3 = Brachytherapy, 4 = External beam, 5 = Intracavitary or interstitial, 8 = Other, B = Radioactive isotopes, X = Unknown / inapplicable		
Radiotherapy description	RADIODESC	Text - no validation		
Imaging code – internal coding system	IMAGINGCODE	Text - no validation		
Description of imaging	IMAGINGDESC	Text - no validation		
Site on body where imaging occurred	IMAGINGSITE	Text - no validation		
List of all systemic anti-cancer therapy drugs	CHEMO_ALL_DRUGS	Text - no validation		
Name or acronym of known drug combinations derived from CHEMO_ALL_DRUGS (e.g. R-CHOP or FEC-T)	CHEMO_DRUG_GROUP	Text - no validation		
Size in millimetres of the diameter of a lesion (histology)	LESIONSIZE	Number or blank		

********** End of sheet **********

IMD income_domain

Introduction

The Income Deprivation Domain measures the proportion of the population experiencing deprivation relating to low income. The definition of low income used includes both those people that are out-of-work, and those that are in work but who have low earnings (and who satisfy the respective means tests).

Data items selection

Please be aware that selecting patient identifiable data items – **the data items marked below in red** – will require you to have patient consent, section 251 approval, or the appropriate statutory regulation covering your organisation. Sensitive patient demographic data, event dates and geographies will be treated as de-personalised, unless anonymisation techniques are applied to render the data anonymised to k-3 anonymity.

Note: only one deprivation quintile per tumour will be provided to lessen the risk of deductive disclosure created by low level geographies, where there have been unique changes to the Income Deprivation Domain overtime. Where the diagnostic period extends across multiple measurement periods (i.e. 2009-2016), a summary variable will be provided describing this measure at the time the tumour was diagnosed.

Data item	Field name	Description of field content	Request field (mark required variables with x)	Justification - detail why the field is necessary for your analysis
Pseudonymised patient ID	PATIENTID	Project specific ID for each person	X	For data management.
Pseudonymised tumour ID	TUMOURID	Project specific ID for each tumour	X	For data management.
Measure of deprivation at small area level made up from the income domain in 2004, quintiles are calculated from populations	IMD2004	1,2,3,4,5		
Measure of deprivation at small area level made up from the income domain in 2007, quintiles are calculated from populations	IMD2007	1,2,3,4,5		
Measure of deprivation at small area level made up from the income domain in 2010, quintiles are calculated from populations	IMD2010	1,2,3,4,5	X	To answer the research questions regarding whether or not targeted therapy utilsation and hormone receptor testing varies by deprivation category.
Measure of deprivation at small area level made up from the income domain in 2015, quintiles are calculated from populations	quintile_2015	1,2,3,4,5	X	To answer the research questions regarding whether or not targeted therapy utilsation and hormone receptor testing varies by deprivation category.

********** End of sheet **********

Appendix 3.2 Continued
SACT

Introduction

Systemic Anti-Cancer Therapy Data Set collects clinical management on patients receiving cancer chemotherapy in or funded by the NHS in England. It relates to all cancer patients, both adult and paediatric, in acute inpatient, daycase, outpatient settings and delivery in the community. It covers chemotherapy treatment for all solid tumour and haematological malignancies and those in clinical trials. The dataset has been designed to collect information on all drug treatments with an anti-cancer effect, in all treatment settings, including traditional cytotoxic chemotherapy and all newer agents.

Data validation is an important aspect of the programme. This is in three stages covering data quality, confirmation of activity levels and an initial clinical analysis. Trusts are required to check the quality, quantity and validity of their data before inclusion for subsequent analysis.

Coverage

Data is available from April 2012 to September 2016 inclusive. The data is linked at a patient level and can be linked to the latest available Cancer Reg- Patient Table

Data item selection

Please be aware that selecting patient identifiable data items – **the data items marked below in red** – will require you to have patient consent, section 251 approval, or the appropriate statutory regulation covering your organisation. Sensitive patient demographic data, event dates and geographies will be treated as de-personalised, unless anonymisation techniques are applied to render the data anonymised to k-3 anonymity. Privacy by design and default principles have also been applied to some of these data to provide a less disclosive alternative. The key explaining how these alternative can be selected is included below:

Key:	
Items in red	Personally identifiable data that require specific lawful permissions to access (as outlined above). In requesting access to these variables, please refer to Section 6 of the 'How to complete the ODR Data Request Form (v4) Guidance' for details of evidence that must accompany your application.
	Derived fields that apply privacy by design and can be used for data minimisation

Data item	**Field name**	**Request field (mark required variables with x)**	**Justification - detail why the field is necessary for your analysis**
Demographics and consultant			
Pseudonymised patient ID	PATIENTID	X	For data management.
Pseudonymised tumour ID	TUMOURID	X	For data management.
NHS number	**NHS_Number**		
NHS number status indicator code	NHS_Number_Status		
Date of birth	**Date_Of_Birth**		
Month of birth	MONTH_DOB		
Year of birth	YEAR_DOB		
Gender code (current)	Gender_Current		
Ethnicity	Ethnicity		
Broad ethnic group	Option to group ethnicities (e.g. white/ non-white/ unknown)		
Postcode	**Postcode**		
Broader geographic area/ IMD quintile	Option to provide geography as deprivation score or aggregate to larger geographic areas such as MSOA or county.		
General medical practice code (patient registration)	GP_Practice_Code		
Consultant code (initiated SACT)	**Consultant_GMC_Code_Clean**		
Consultant code (pseudonymised)	Consultant_GMC_Code (pseudonymised)		
Care professional main speciality code (start SACT)	Consultant_Speciality_Code		
Organisation code	Organisation_Code_of_Provider		
Organisation code (pseudonymised)	Organisation_Code_of_Provider (pseudonymised)		

Appendix 3.2 Continued

	Clinical status			
	Primary diagnosis (on SACT initiation)	Primary_Diagnosis		
*	Morphology (ICD-O on SACT initiation)	Morphology_clean		
	Pre- treatment (final) TNM stage	Stage_at_Start		
	Programme and regimen			
	SACT programme number	Programme_Number	X	For data management.
	Anti-cancer regimen number	Regimen_Number	X	For descriptive analyses.
**	Drug treatment intent	Intent_of_Treatment	X	For descriptive and adjusted analyses (propensity score matching).
***	Regimen analysis grouping	Analysis_Group	X	For descriptive and adjusted analyses to group data by drug name/drug class.
***	Regimen grouping (benchmark reports)	Benchmark_Group	X	For descriptive and adjusted analyses to group data by drug name/drug class.
	Patient's height (metres (m))	Height_At_Start_of_Regimen		
	Patient's weight (kilograms (kg))	Weight_At_Start_of_Regimen		
	Performance Status (Adult)	Performance_Status_at_Start_of_Regimen_Clean	X	For descriptive and adjusted analyses.
	Performance Status (Young Person)	Performance_Status_at_Start_of_Regimen_Clean		
****	Co-morbidity adjustment indicator	Comorbidity_Adjustment		
	Decision to treat date (Drug regimen)	Date_Decision_To_Treat		
	Month of decision to treat (Drug regimen)	Month of decision to treat		
	Year of decision to treat (Drug regimen)	Year of decision to treat		
	Days from another event to decision to treat date	Option to provide number of days from another event to the date of the decision to treat (e.g. days from diagnosis to date of decision to treat)		
	Start date (Drug regimen)	Start_Date_of_Regimen		
	Month of start date for drug regimen	Month of start date of drug regimen	X	For descriptive and adjusted analyses.
	Year of start date for drug regimen	Year of start date of drug regimen	X	For descriptive and adjusted analyses.
	Days from another event to drug regimen start date	Option to provide number of days from another event to the start date of the drug regimen (e.g. days from date of decision to treat to start date of regimen)		
	Clinical trial indicator	Clinical_Trial		
	Chemo-radiation indicator	Chemo_Radiation		
	Number of planned systemic anti-cancer therapy cycles	Number_of_Cycles_Planned		
	Cycle			
	Cycle identifier	Cycle_Number		
	Start date (Cycle)	Start_Date_of_Cycle		
	Month of start date of cycle	Month of start date of cycle		
	Year of start date of cycle	Year of start date of cycle		
	Days from another event to start date of cycle	Option to provide number of days from another event to the start date of the cycle (e.g. days from diagnosis to start date of cycle)		
	Patient's Weight (Kilograms (kg))	Weight_At_Start_Of_Cycle		
*****	Performance Status (Adult)	Performance_Status_At_Start_Of_Cycle_Clean		

Notes from Sarah Lawton (PHE, April 2019):

*SACT morphology data are not well recorded or accurate – request from NCRD instead.

**Treatment intent records palliative/curative descriptor groups – this field is not necessarily accurately recorded.

***Regimen name or drug name is rather messy data, though ODR has enlisted pharmacists to help clean up this data field. Recommend consulting with a specialist if analysing this data field.

****Comorbidity adjustment indicator in SACT is a measure of how significant comorbidity was as a factor in treatment decisions. Request Charlson Comorbidity Index from the NCRD if require number of comorbidities that a patient has.

*****Performance status is at the start of the cycle so can vary.

Appendix 3.2 Continued

Performance Status (Young Person)	Performance_Status_At_Start_Of_Cycle_Clean		
Primary procedure (OPCS)	OPCS_Procurement_Code		
Drug details			
Drug analysis grouping	Drug_Group	X	To aid targeted therapy group classification. All treatments bar non-chemotherapy supportive dru
Actual dose	Actual_Dose_Per_Administration	X	For descriptive and adjusted analyses.
SACT drug route of administration	Administration_Route	X	For descriptive and adjusted analyses.
SACT administration date	Administration_Date	X	For temporal analyses.
Organisation code (provider)	Organisation_Code_of_Drug_Provider		
Pseudonymised organisation code (provider)	Organisation_Code_of_Drug_Provider (pseudonymised)		
Primary procedure (OPCS)	OPCS_Delivery_Code		
Outcome			
Start date (Final therapy)	Date_of_Final_Treatment		
Month of final therapy	Month of final therapy	X	For survival analyses in order to determine treatment duration.
Year of final therapy	Year of final therapy	X	For survival analyses in order to determine treatment duration.
Days from another event to start date of final therapy	Option to provide number of days from another event to the start date of the final therapy (e.g. days from diagnosis to start date of final therapy)		
Regimen modification indicator (dose reduction)	Regimen_Modification_Dose_Reduction	X	To establish the cohort numbers requiring intervention.
Regimen modification indicator (time delay)	Regimen_Modification_Time_Delay	X	To establish the cohort numbers requiring intervention.
Regimen modification indicator (days reduced)	Regimen_Modification_Stopped_Early	X	To establish the cohort numbers requiring intervention.
Planned treatment change reason	Regimen_Outcome_Summary	X	For adjusted analyses.

********** **End of sheet** **********

Additional notes from Sarah Lawton (PHE, April 2019):

- Outcome information likely poorly recorded, not a complete record as data field relies on prescribing consultant to have closed the record down which does not always occur. Though there is some information which could prove useful for determining reasons for stopping a drug.
- Do not request the final treatment data information as this is an automatic system data recorded for the last cycle which may not be perfect.
- There are potentially some issues with data recording in Trusts who have only submitted data on the 1st drug cycle. It may therefore not be possible to determine if a patient only ever had one cycle or went on to receive further ones which were not recorded electronically. More problematic in Trusts using paper-based drug recording.
- Suggests accessing for area-level descriptors at a regional as opposed to Trust level.

Appendix 3.3 Variables requested from the NCRD for Chapter 3

Variable	Derived from ODR variable (by ODR)	Reason for Inclusion	Analyses Used in
Pseudonymised Patient ID	PATIENTID	For data management to link files.	None
Pseudonymised Tumour ID	TUMOURID	For data management to link files.	None
Site of Neoplasm	SITE_ICD10_02	To differentiate breast and lung patients (as well as those with both tumour types) into the appropriate analyses.	None
Cancer Morphology	MORPH__CD10_02	Used to classify lung cancer histology and define the denominator population for the cohort of interest.	None
Tumour Grade	GRADE	Tumour grade may affect utilisation and will be included as a confounder.	Logistic regression analyses
Deprivation Status (Index of Multiple Deprivation)	IMD2010 Quintile_2015	Variable of focus for this thesis. IMD will be the measure of deprivation used.	Logistic regression analyses
Sex	SEX	Sex may affect utilisation and will be included as a confounder.	Logistic regression analyses
Ethnic Group	ETHNICITYNAME	Ethnicity may affect utilisation and will be included as a confounder.	Logistic regression analyses
Age at Diagnosis in 5 Year Age Bands	FIVEYEARAGEBAND	Age may affect utilisation and will be included as a confounder. This variable was requested in 5 year time bands in case actual day, month and year of birth was not provided.	Logistic regression analyses
Diagnosis Date	DIAGNOSISDATEBEST DIAGNOSISDATEMONTH DIAGNOSISDATEYEAR	Used to refine cohorts of interest by time for typical treatment periods. Diagnosis year used in analyses.	Temporal trends in utilisation. Included in logistic regression analyses.

Appendix 3.3 Continued

Variable	Derived from ODR variable (by ODR)	Reason for Inclusion	Analyses Used in
Pre-treatment TNM	T_IMG N_IMG M_ING T_BEST N_BEST M_BEST	Requested for descriptive and adjusted analyses but final modelling used stage best variable instead.	None
Stage (Best)	STAGE_BEST	Stage may affect utilisation and will be included as a confounder.	Logistic regression analyses
Count of Evert Tumour Assigned to that Patient ID	BIGTUMOURCOUNT		Logistic regression analyses
Oestrogen Receptor Status	ER_STATUS ER_SCORE	ER Status may affect utilisation and will be included as a confounder.	Logistic regression analyses
Progesterone Receptor Status	PR_STATUS PR_SCORE	PR Status was requested in case this was of interest for drug utilisation. However, upon speaking with a breast oncologist, this variable was dropped from the analyses as it is less useful for treatment decisions in HER2+ breast cancer patients.	None
HER2 Tumours Status	HER2_STATUS	Used to refine the cohort of breast cancer patients and to define the denominator population for the breast analysis.	None
Government Region for Patient Residence	GOR_NAME	We applied for trust code, Clinical Commissioning Group code, Primary Care Trust code and Local Authority Unitary Codes too as we were unsure at what level detail we would be given such geographical information.	Logistic regression analyses

Appendix 3.3 Continued

Variable	Derived from ODR variable (by ODR)	Reason for Inclusion	Analyses Used in
Surgery Within 6 Months Flag	SG_WITHIN_SIX_MONTHS_FLAG	Included as a comparator for utilisation with deprivation and previous studies have shown that these two variables are linked.	Logistic regression analyses
Chemotherapy Within 6 Months Flag	CT_WITHIN_SIX_MONTHS_FLAG	Included as a comparator for utilisation with deprivation as previous studies have shown that these two variables are associated.	Logistic regression analyses
Radiotherapy Within 6 Months Flag	RT_WITHIN_SIX_MONTHS_FLAG	Requested for potential inclusion as a compactor and whilst we did explore some descriptive statistics with this variable, these were not used in the thesis results.	None
Route to Diagnosis Code	ROUTE_CODE FINAL_ROUTE	Requested for potential inclusion in regression analyses but later omitted on account of the amount of missing data.	None
Multidisciplinary Team Indicator	MDT_IND	MDT discussion may affect utilisation and will be included as a confounder.	Logistic regression analyses
Vital Status of the Patient	VITALSTATUS	Requested for survival analyses.	None
Days from Another Event to Date to Death		Requested for survival analyses.	None
Death Notification	DEATHCAUSECODE	Requested for survival analyses.	None
Diagnosis Death Certificate Only	DCO	Requested for survival analyses.	None
All HES Diagnosis Codes	Diag3_3n	Requested to identify specific comorbidities (12 months prior to and post discharge).	None
All HES Operative Procedures Codes	Opertn_nn	Requested to expand on details for surgery receipt.	None

DCO: Death certificate only: ER: Oestrogen receptor; HER2: Human epidermal growth receptor 2; HES: Hospital episode statistics; ICD: International Classification of Disease: ID: Identifier; IMD: Index of multiple deprivation; MDT: Multidisciplinary team; ODR: Office for Data Release; PR: Progesterone receptor; TNM: Tumour, nodes, metastases

Appendix 3.4 Variables requested from the SACT dataset for Chapter 3

Variable	Derived from ODR variable (by ODR)	Reason for Inclusion	Analyses Used in
Pseudonymised Patient ID	PATIENTID	For data management to link files	None
Pseudonymised Tumour ID	TUMOURID	For data management to link files	None
SACT Programme Number	Programme_Number	For data management	None
Anti-Cancer Regimen Number	Regimen_Number	For data management	None
Drug Treatment Intent	Intent_of_Treatment	For data analysis	None
Regimen Analysis Grouping	Analysis_Group	Used to create a binary variable for SACT treatment receipt along with benchmark and drug groups.	Logistic regression analyses
Regimen Grouping (Benchmark Reports)	Benchmark_Group	Used to create a binary variable for SACT treatment receipt along with analysis and drug groups.	Logistic regression analyses
Performance Status	Performance_Status_at_Start_of_Regimen_Clean	Include as a potential confounder but too much missing information to do so.	None
Drug Regimen Start Date	Month of start date of drug regimen Year of start date of drug regimen	Used to determine the time frame for drug use for the analysis.	Data cleaning
Drug Analysis Grouping	Drug_Group	Used to create a binary variable for SACT treatment receipt along with benchmark and regimen groups.	Logistic regression analyses

Appendix 3.4 Continued

Variable	Derived from ODR variable (by ODR)	Reason for Inclusion	Analyses Used in
Actual Dose	Actual_Dose_Per_Administration	Requested in case this variable helped with drug identification of interest. Realised that this level of detail was not required for the thesis analyses.	None
SACT Drug Administration Route	Administration_Route	Requested in case this variable helped with drug identification of interest. Realised that this level of detail was not required for the thesis analyses.	None
SACT Administration Date	Administration_Date	Variable was obtained to potentially help refine the denominator population of interest by time from drug date.	Data cleaning
Final Therapy Date	Month of final therapy Year of final therapy	Variable was requested in case this was of use for time analyses but only start date was needed.	None
Regimen Modifier (Dose Reduction, Time Delay, Stopped Early, Planned Treatment Change)	Regimen_Modification_Dose_Reduction Regimen_Modification_Time_Delay Regimen_Modification_St Regimen_Outcome_Summary	Requested in case was of interest for analyses but felt this level of detail did not add to an analysis on treatment utilisation.	None

ODR: Office for Data Release; ID: Identifier; SACT: Systemic Anti-Cancer therapy.

Appendix 3.5 Trastuzumab SACT references used for coding as trastuzumab receipt for Chapter 3

SACT Drug Classification	SACT Analysis Group Listings	SACT Benchmark Group Listings	SACT Drug Group Listings
TRASTUZUMAB Anti-HER2 Targeted Therapy (MAB)	CAPECITABINE + **TRASTUZUMAB** CARBOPLATIN + DOCETAXEL + PERTUZUMAB + **TRASTUZUMAB** CARBO + DOCETAXEL + **TRASTUZUMAB** CARBOPLATIN + DOCETAXEL + **TRASTUZUMAB** CARBOPLATIN + PACLITAXEL + **TRASTUZUMAB** CARBOPLATIN + **TRASTUZUMAB** CYCLO + DOCETAXEL + **TRASTUZUMAB** CYCLOPHOSPHAMIDE + DOCETAXEL + DOXORUBICIN + PERTUZUMAB + **TRASTZUMAB** CYCLOPHOSPHAMIDE + DOCETAXEL + DOXORUBICIN + **TRASTUZUMAB** CYCLOPHOSPHAMIDE + DOCETAXEL + PERTUZUMAB + **TRASTUZUMAB** CYCLOPHOSPHAMIDE + DOCETAXEL + **TRASTUZUMAB** CYCLOPHOSPHAMIDE + DOXORUBICIN + PERTUZUMAB + **TRASTUZUMAB** DOCETAXEL + PERTUZUMAB + **TRASTUZUMAB** DOCETAXEL + **TRASTUZUMAB** EC + DOCETAXEL + **TRASTUZUMAB** FEC + DOCETAXEL + PERTUZUMAB + **TRASTUZUMAB** FEC + DOCETAXEL + **TRASTUZUMAB** FEC 100 + DOCETAXEL + **TRASTUZUMAB** FEC + PACLITAXEL + **TRASTUZUMAB** FEC + **TRASTUZUMAB** FEC 60 OR 75 + **TRASTUZUMAB** FEC 100 + **TRASTUZUMAB** LAPATINIB + **TRASTUZUMAB** NAB-PACLITAXEL + **TRASTUZUMAB** PACLITAXEL + PERTUZUMAB + **TRASTUZUMAB** PACLITAXEL + **TRASTUZUMAB** PERTUZUMAB + **TRASTUZUMAB** TAC + **TRASTUZUMAB** **T**CH TRASTUZUMAB **TRASTUZUMAB** + VINORELBINE APHINITY TRIAL BERENICE TRIAL EPHOS-B TRIAL METAPHER TRIAL PERSEPHONE TRIAL ROSCO TRIAL SAFEHER TRIAL SOLD TRIAL	CAPECITABINE + **TRASTUZUMAB** CARBO + DOCETAXEL + **TRASTUZUMAB** CARBOPLATIN + DOCETAXEL + PERTUZUMAB + **TRASTUZUMAB** CARBOPLATIN + DOCETAXEL + **TRASTUZUMAB** CARBOPLATIN + PACLITAXEL + **TRASTUZUMAB** CARBOPLATIN + **TRASTUZUMAB** CYCLO + DOCETAXEL + **TRASTUZUMAB** CYCLOPHOSPHAMIDE + DOCETAXEL + DOXORUBICIN + PERTUZUMAB + **TRASTUZUMAB** CYCLOPHOSPHAMIDE + DOCETAXEL + DOXORUBICIN + **TRASTUZUMAB** CYCLOPHOSPHAMIDE + DOCETAXEL + PERTUZUMAB + **TRASTUZUMAB** CYCLOPHOSPHAMIDE + DOCETAXEL + **TRASTUZUMAB** CYCLOPHOSPHAMIDE + DOXORUBICIN + PERTUZUMAB + **TRASTUZUMAB** DOCETAXEL + PERTUZUMAB + **TRASTUZUMAB** DOCETAXEL + **TRASTUZUMAB** EC + DOCETAXEL + **TRASTUZUMAB** FEC + DOCETAXEL + PERTUZUMAB **+ TRASTUZUMAB** FEC + DOCETAXEL + **TRASTUZUMAB** FEC + PACLITAXEL + **TRASTUZUMAB** FEC + **TRASTUZUMAB** LAPATINIB + **TRASTUZUMAB** NAB-PACLITAXEL + **TRASTUZUMAB** PACLITAXEL + PERTUZUMAB + **TRASTUZUMAB** PACLITAXEL + **TRASTUZUMAB** PERTUZUMAB + **TRASTUZUMAB** TAC + **TRASTUZUMAB** **T**CH TRASTUZUMAB **TRASTUZUMAB** + VINORELBINE APHINITY TRIAL BERENICE TRIAL EPHOS-B TRIAL METAPHER TRIAL PERSEPHONE TRIAL ROSCO TRIAL SAFEHER TRIAL SOLD TRIAL	**TRASTUZUMAB** **TRASTUZUMAB** (HERCEPTIN) **TRASTUZUMAB** BIOSIMILAR (HERZUMA) **TRASTUZUMAB** BIOSIMILAR (ONTRUZANT)

Appendix 3.5 Continued

*Trials listed in table only if it is certain that all patients received that drug.
Bold text indicates the reference to trastuzumab within the drug regimen.

HER2 Trial Information
APHINITY Trial – Group 1 receives chemotherapy, trastuzumab and pertuzumab. Group 2 receive chemotherapy, trastuzumab and a placebo.
BERENICE Trial – Group 1 receives doxorubicin and cyclophosphamide then paclitaxel plus trastuzuamb and pertuzumab. Group 2 receives fluorouracil, epirubicin, cyclophosphamide, then docetaxel with trastuzumab and pertuzumab.
EPHOS-B Trial – Everyone receives chemotherapy and trastuzumab after surgery. Group 1 have surgery, chemotherapy then trastuzumab. Group 2 have trastuzumab, before and after surgery, followed by chemotherapy then trastuzumab. Group 3 have lapatinib and trastuzumab before and after surgery followed by chemotherapy then trastuzumab
KATHERINE Trial – Patients received either trastuzumab emtansine or trastuzumab.
MetaPHER Trial – Everyone receives docetaxel after trastuzumab and pertuzumab.
PERSEPHONE Trial – Everyone receives trastuzumab (either for 6 or 12 months) with or after chemotherapy.
ROSCO Trial – Either receive docetaxel and cyclophosphamide or FEC. If HER2+ also receive trastuzumab.
SafeHER Trial – Receive SC Trastuzumab via assisted administration conventional syringe and needle/vial formulation or with assisted or self-administration using a single-se injection device.
SOLD Trial – Everyone had docetaxel and trastuzumab at the same time to start then FEC. Some patients then received more trastuzumab and some didn't.

Abbreviations: EC: Epirubicin & Cyclophosphamide; FEC: Fluorouracil, Epirubicin & Cyclophosphamide; MAB: Monoclonal Antibody; SACT: Systemic Anti-Cancer Therapy; TAC: Docetaxel, Doxorubicin & Cyclophosphamide; TCH: Docetaxel, Carboplatin & Trastuzumab.

Appendix 3.6 Demographic and clinical characteristics of a female cohort with a first invasive primary HER2+ breast cancer diagnosed between 01/01/2012 – 31/12/2017 who had a SACT record (n = 25,599), had a SACT record in the time range (n = 21,881) and who did not have a SACT record (n = 14,580) for Chapter 3

	Had a SACT Record n =25,599	Had a SACT Record in the Time Range[1] n = 21,881	Did not have a SACT Record n = 14,580
		Number (%)	
IMD[2]			
1 (Least Deprived)	5,687 (22.22)	4,960 (22.67)	3,585 (24.59)
2	5,644 (22.05)	4,849 (22.16)	3,323 (22.79)
3	5,174 (20.21)	4,335 (19.81)	2,953 (20.25)
4	4,770 (18.63)	4,086 (18.67)	2,605 (17.87)
5 (Most Deprived)	4,324 (16.89)	3,651 (16.69)	2,114 (14.50)
Age (Years)			
<40	2,231 (8.72)	2,019 (9.23)	385 (2.64)
40 - 49	5,665 (22.13)	5,002 (22.86)	1,489 (10.21)
50 – 59	7,180 (28.05)	6,337 (28.96)	2,598 (17.82)
60 – 69	6,358 (24.84)	5,445 (24.88)	3,236 (22.19)
70 – 79	3,224 (12.59)	2,573 (11.76)	3,251 (22.30)
80+	941 (3.68)	505 (2.31)	3,621 (24.84)
Diagnosis Year			
2012	2,885 (11.27)	1,840 (8.41)	1,831 (12.56)
2013	4,024 (15.72)	3,025 (13.82)	1,721 (11.80)
2014	4,111 (16.06)	3,434 (15.69)	2,108 (14.56)
2015	4,424 (17.28)	3,963 (18.11)	2,556 (17.53)
2016	4,956 (19.36)	4,632 (21.17)	2,946 (20.21)
2017	5,199 (20.31)	4,987 (22.79)	3,418 (23.44)
Ethnicity			
White	22,645 (88.46)	19,294 (88.18)	12,631 (86.63)
Other Ethnic Group[3]	2,205 (8.61)	1,939 (8.86)	885 (6.07)
Missing/unknown[4]	749 (2.93)	648 (2.96)	1,064 (7.30)
Rural/Urban Indicator			
Rural Village, Hamlet & Isolated Dwellings	2,760 (10.78)	2,366 (10.81)	1,614 (11.07)
Rural Town & Fringe	2,749 (10.74)	2,340 (10.69)	1,546 (10.60)
Urban City & Town	11,510 (44.96)	9,685 (44.26)	6,755 (46.33)
Urban Conurbation	8,580 (33.52)	7,490 (34.23)	4,665 (32.00)
Government Region			
North West	4,067 (15.89)	3,691 (16.87)	2,242 (15.38)
North East	1,797 (7.02)	1,387 (6.34)	735 (5.04)
West Midlands	2,730 (10.66)	2,124 (9.71)	1,572 (10.78)
Yorkshire & the Humber	2,417 (9.44)	2,129 (9.73)	1,143 (7.84)
East Midlands	2,434 (9.51)	2,265 (10.35)	1,221 (8.37)
East of England	3,265 (12.75)	2,709 (12.38)	1,923 (13.19)
South East	3,832 (14.97)	3,259 (14.89)	2,651 (18.18)
South West	2,704 (10.56)	2,225 (10.17)	1,645 (11.28)
London	2,353 (9.19)	2,092 (9.56)	1,448 (9.93)

Appendix 3.6 Continued

	Had a SACT Record n =25,599	Had a SACT Record in the Time Range[1] n = 21,881	Did not have a SACT Record n = 14,580
	Number (%)		
Stage			
I	6,739 (26.33)	5,662 (25.88)	6,355 (43.59)
II	11,664 (45.56)	10,188 (46.56)	4,999 (34.29)
III	3,881 (15.16)	3,395 (15.52)	978 (6.71)
IV	1,636 (6.39)	1,329 (6.07)	733 (5.03)
Unknown[5]	1,679 (6.56)	1,307 (5.97)	1,515 (10.39)
Grade			
Well differentiated (Low Grade)	599 (2.34)	427 (1.95)	1,578 (10.82)
Moderately Differentiated	9,761 (38.13)	8,219 (37.56)	7,630 (52.33)
Poorly Differentiated	14,740 (57.58)	12,827 (58.62)	5,024 (34.46)
Other[6]	499 (1.95)	408 (1.86)	348 (2.39)
Big Tumour Count			
1	22,724 (88.77)	19,625 (89.69)	12,579 (86.28)
>1	2,875 (11.23)	2,256 (10.31)	2,001 (13.72)
ER Status			
Positive	15,241 (59.54)	12,651 (57.82)	10,230 (70.16)
Negative	6,399 (25.00)	5,586 (25.53)	2,202 (15.10)
Unknown	3,959 (15.47)	3,644 (16.65)	2,148 (14.73)
CCM (Between 78 to 6 Months Prior to Diagnosis)			
0	21,917 (85.62)	18,922 (86.48)	10,982 (75.32)
1-2	3,216 (12.56)	2,644 (12.08)	2,686 (18.42)
3+	466 (1.82)	315 (1.44)	912 (6.26)
Discussed at MDT			
Yes	18,711 (73.09)	15,875 (72.55)	9,072 (62.22)
No	3,820 (14.92)	3,340 (15.26)	2,417 (16.58)
Missing	3,068 (11.98)	2,666 (12.18)	3,091 (21.20)
SACT Record			
Yes (In Time Range)[1]	21,881(85.48)	21,881 (100.00)	----------
Yes (Not in Time Range)[1]	14,3,718 (14.52)	----------	----------
Treatment			
Utilised Chemotherapy	23,161 (90.48)	20,978 (95.87)	1,283 (8.80)
Utilised Surgery	2,438 (9.52)	16,646 (76.08)	10,716 (73.50)
Utilised Trastuzumab	17,674 (69.04)	17,674 (80.77)	----------

[1]In time range refers to a SACT record up to 56 days before and 1 year post diagnosis.
[2]For diagnosis year 2012, IMD_2010 was used and for diagnosis years 2013 – 2107, IMD_2015 was used.
[3]Other refers to Asian/British Asian, Black/African/Caribbean/Black British, mixed/multiple ethnic groups and other ethnic groups.
[4]Missing/unknown refers to unknown and missing ethnicity classifications.
[5]Unknown staging refers to missing and unstageable tumours.
[6]Other refers to undifferentiated or anaplastic, undetermined, and missing tumour grades.
CCM, Charlson Comorbidity Index; ER, Oestrogen receptor status; IMD: Index of multiple deprivation; MDT: Multidisciplinary team: SACT: Systemic Anti-Cancer Therapy.

Appendix 3.7 Demographic and clinical characteristics by deprivation category for a female cohort with HER2+ breast cancer diagnosed between 01/01/2012 – 31/12/2017 (n = 40,179) for Chapter 3

	IMD 1[1] (Least Deprived) N (%)	IMD 2 N (%)	IMD 3 N (%)	IMD 4 N (%)	IMD 5 (Most Deprived) N (%)	P Value[2]
Age (Years)						<0.001
<40	493 (5.32)	502 (5.60)	548 (6.74)	539 (7.31)	534 (8.29)	
40 – 49	1,653 (17.83)	1,526 (17.02)	1,393 (17.14)	1,362 (18.47)	1,220 (18.95)	
50 – 59	2,241 (24.17)	2,204 (24.58)	1,939 (23.86)	1,810 (24.54)	1,584 (24.60)	
60 – 69	2,283 (24.62)	2,249 (25.08)	1,961 (24.13)	1,683 (22.82)	1,418 (22.03)	
70 – 79	1,586 (17.11)	1,478 (16.48)	1,294 (15.92)	1,146 (15.54)	971 (15.08)	
80+	1,016 (10.96)	1,008 (11.24)	992 (12.21)	835 (11.32)	711 (11.04)	
Diagnosis Year						0.676
2012	1,068 (11.52)	1,055 (11.77)	993 (12.22)	860 (11.66)	740 (11.49)	
2013	1,355 (14.61)	1,265 (14.11)	1,155 (14.21)	1,030 (13.97)	940 (14.60)	
2014	1,468 (15.83)	1,416 (15.79)	1,205 (14.83)	1,130 (15.32)	1,000 (15.53)	
2015	1,577 (17.01)	1,563 (17.43)	1,414 (17.40)	1,267 (17.18)	1,159 (18.00)	
2016	1,865 (20.11)	1,771 (19.75)	1,588 (19.54)	1,449 (19.65)	1,229 (19.09)	
2017	1,939 (20.91)	1,897 (21.16)	1,772 (21.80)	1,639 (22.22)	1,370 (21.28)	
Ethnicity						<0.001
White	8,353 (90.09)	8,089 (90.21)	7,243 (89.12)	6,334 (85.88)	5,257 (81.66)	
Other Ethnic Group[3]	348 (3.75)	400 (4.46)	550 (6.77)	793 (10.75)	999 (15.52)	
Missing/unknown[4]	571 (6.16)	478 (5.33)	334 (4.11)	248 (3.36)	182 (2.83)	
Rural/Urban Indicator						<0.001
Rural Village, Hamlet & Isolated Dwellings	1,639 (17.68)	1,738 (19.38)	815 (10.03)	151 (2.05)	31 (0.48)	
Rural Town & Fringe	1,184 (12.77)	1,268 (14.14)	1,043 (12.83)	635 (8.61)	165 (2.56)	
Urban City & Town	4,436 (47.84)	3,861 (43.06)	3,861 (47.51)	3,632 (49.25)	2,475 (38.44)	
Urban Conurbation	2,013 (21.71)	2,100 (23.42)	2,408 (29.63)	2,957 (40.09)	3,767 (58.51)	

Appendix 3.7 Continued

	IMD 1[1] (Least Deprived) N (%)	IMD 2 N (%)	IMD 3 N (%)	IMD 4 N (%)	IMD 5 (Most Deprived) N (%)	P Value[2]
Government Region						<0.001
North West	1,254 (13.52)	1,308 (14.59)	1,063 (13.08)	1,108 (15.02)	1,576 (24.48)	
North East	460 (4.96)	413 (4.61)	411 (5.06)	547 (7.42)	701 (10.89)	
West Midlands	819 (8.83)	915 (10.20)	826 (10.16)	754 (10.22)	988 (15.35)	
Yorkshire & the Humber	810 (8.74)	799 (8.91)	654 (8.05)	599 (8.12)	698 (10.84)	
East Midlands	832 (8.97)	874 (9.75)	768 (9.45)	700 (9.49)	481 (7.47)	
East of England	1,291 (13.92)	1,328 (14.81)	1,230 (15.13)	932 (12.64)	407 (6.32)	
South East	2,281 (24.60)	1,565 (17.45)	1,332 (16.39)	909 (12.33)	396 (6.15)	
South West	1,050 (11.32)	1,244 (13.87)	1,066 (13.12)	721 (9.78)	268 (4.16)	
London	475 (5.12)	521 (5.81)	777 (9.56)	1,105 (14.98)	923 (14.34)	
Stage						<0.001
I	3,154 (34.02)	3,106 (34.64)	2,684 (33.03)	2,279 (30.90)	1,871 (29.06)	
II	3,805 (41.04)	3,657 (40.78)	3,390 (41.71)	3,090 (41.90)	2,721 (42.26)	
III	1,011 (10.90)	1,023 (11.41)	950 (11.69)	930 (12.61)	945 (14.68)	
IV	484 (5.22)	481 (5.36)	445 (5.48)	515 (6.98)	444 (6.90)	
Unknown[5]	818 (8.82)	700 (7.81)	658 (8.10)	561 (7.61)	457 (7.10)	
Grade						0.001
Well Differentiated (Low Grade)	471 (5.08)	493 (5.50)	484 (5.96)	397 (5.38)	332 (5.16)	
Moderately Differentiated	4,026 (43.42)	3,928 (43.81)	3,517 (43.28)	3,090 (41.90)	2,830 (43.96)	
Poorly Differentiated	4,538 (48.94)	4,372 (48.76)	3,941 (48.49)	3,752 (50.87)	3,161 (49.10)	
Other[6]	237 (2.56)	174 (1.94)	185 (2.28)	136 (1.84)	115 (1.79)	
Big Tumour Count						0.160
1	8,100 (87.36)	7,890 (87.99)	7,118 (87.58)	6,489 (87.99)	5,706 (88.63)	
>1	1,172 (12.64)	1,077 (12.01)	1,009 (12.42)	886 (12.01)	732 (11.37)	

Appendix 3.7 Continued

	IMD 1[1] (Least Deprived) N (%)	IMD 2 N (%)	IMD 3 N (%)	IMD 4 N (%)	IMD 5 (Most Deprived) N (%)	P Value[2]
ER Status						<0.001
Positive	5,732 (61.82)	5,770 (64.35)	5,231 (64.37)	4,688 (63.57)	4,050 (62.91)	
Negative	1,924 (20.75)	1,835 (20.46)	1,766 (21.73)	1,644 (22.29)	1,432 (22.24)	
Unknown	1,616 (17.43)	1,362 (15.19)	1,130 (13.90)	1,043 (14.14)	956 (14.85)	
CCM (Between 78 to 6 Months Prior to Diagnosis)						<0.001
0	7,898 (85.18)	7,541 (84.10)	6,656 (81.90)	5,886 (79.81)	4,918 (76.39)	
1 - 2	1,157 (12.48)	1,175 (13.10)	1,195 (14.70)	1,194 (16.19)	1,181 (18.34)	
3+	217 (2.34)	251 (2.80)	276 (3.40)	295 (4.00)	339 (5.27)	
Discussed at MDT						<0.001
Yes	6,254 (67.45)	6,269 (69.91)	5,693 (70.05)	5,084 (68.94)	4,483 (69.63)	
No	1,312 (14.15)	1,313 (14.64)	1,266 (15.58)	1,243 (16.85)	1,103 (17.13)	
Missing	1,706 (18.40)	1,385 (15.45)	1,168 (14.37)	1,048 (14.21)	852 (13.23)	
SACT Record						<0.001
Yes (In Time Range)[7]	4,960 (53.49)	4,849 (54.08)	4,335 (53.34)	4,086 (55.40)	3,651 (56.71)	
Yes (Not in Time Range)[7]	727 (7.84)	795 (8.87)	839 (10.32)	684 (9.27)	673 (10.45)	
No	3,585 (38.66)	3,323 (37.06)	2,953 (36.34)	2,605 (35.32)	2,114 (32.84)	
Treatment						
Received Chemotherapy	5,494 (59.25)	5,387 (60.08)	4,949 (60.90)	4,517 (61.25)	4,097 (63.64)	<0.001
Received Surgery	7,075 (76.31)	6,948 (77.48)	6,109 (75.17)	5,397 (73.18)	4,595 (71.37)	<0.001
Received Trastuzumab	4,088 (44.09)	3,959 (44.15)	3,456 (42.52)	3,266 (44.28)	2,905 (45.12)	0.029

[1]For diagnosis year 2012, IMD_2010 was used and for diagnosis years 2013 - 2017, IMD_2015 was used; [2]Chi-Square P Values; [3]Missing/unknown refers to unknown and missing ethnicity classifications; [4]Other refers to Asian/British Asian, Black/African/Caribbean/Black British, mixed/multiple ethnic groups and other ethnic group; [5]Unknown staging refers to missing and unstageable tumours; [6]Other refers to undifferentiated or anaplastic, undetermined and missing tumour grades; [7]From 56 days prior to and up to 365 days post diagnosis. CCM: Charlson Comorbidity Index; ER: Oestrogen receptor status: IMD: Index of multiple deprivation; MDT: Multi-disciplinary team; SACT: Systemic Anti-Cancer Therapy.

Appendix 3.8 Sensitivity analyses for trastuzumab utilisation for Chapter 3

				Unadjusted			Adjusted		
	Number (%) Utilising Trastuzumab	Number (%) Not Utilising Trastuzumab	P Value[1]	OR	95% CI	P Value[2]	OR	95% CI	P Value[2]
Full Analytical Cohort Original Model for Comparison (n = 40,179)									
IMD[3]	n = 17,674 (43.99)	n = 22,505 (56.01)	0.029			**0.0285**			**0.0396**
1 (Least Deprived)	4,088 (44.09)	5,184 (55.91)		1.00	----- -----	-------	1.00	----- -----	-------
2	3,959 (44.15)	5,008 (55.85)		1.00	0.95 – 1.06	0.934	0.99	0.93 – 1.06	0.846
3	3,456 (42.52)	4,671 (57.48)		0.94	0.88 – 1.00	0.038	0.92	0.86 – 0.99	0.021
4	3,266 (44.28)	4,109 (55.72)		1.01	0.95 – 1.07	0.801	0.93	0.86 – 1.00	0.040
5 (Most Deprived)	2,905 (45.12)	3,533 (54.88)		1.04	0.98 – 1.11	0.200	0.92	0.85 – 0.99	0.036
Diagnosis Date Post Mandated SACT Submission 01/04/2014 – 31/12/2017 (n = 28,146) Sensitivity Analysis 1									
IMD[3]	n = 12,835 (45.60)	n = 15,311 (54.40)	0.028			**0.0279**			**0.1356**
1 (Least Deprived)	2,921 (45.04)	3,565 (54.96)		1.00	----- -----	-------	1.00	----- -----	-------
2	2,902 (46.04)	3,401 (53.96)		1.04	0.97 – 1.12	0.253	1.06	0.97 – 1.14	0.191
3	2,497 (44.08)	3,168 (55.92)		0.96	0.90 – 1.03	0.289	0.96	0.88 – 1.04	0.343
4	2,409 (46.26)	2,798 (53.74)		1.05	0.98 – 1.13	0.185	0.98	0.90 – 1.07	0.640
5 (Most Deprived)	2,106 (46.96)	2,379 (53.04)		1.08	1.00 – 1.17	0.047	0.95	0.86 – 1.04	0.248
Positive HER2+ Status Definition (n = 27,712) Sensitivity Analysis 2									
IMD[3]	n = 15,944 (57.53)	n = 11,768 (42.47)	0.199			**0.1987**			**0.0730**
1 (Least Deprived)	3,719 (57.83)	2,712 (42.17)		1.00	----- -----	-------	1.00	----- -----	-------
2	3,561 (57.30)	2,654 (42.70)		0.98	0.91 – 1.05	0.545	0.94	0.87 – 1.02	0.162
3	3,136 (56.33)	2,431 (43.67)		0.94	0.87 – 1.01	0.098	0.91	0.84 – 0.99	0.033
4	2,940 (57.82)	2,145 (42.18)		1.00	0.93 – 1.08	0.990	0.90	0.82 – 0.98	0.019
5 (Most Deprived)	2,588 (58.63)	1,826 (41.37)		1.03	0.96 – 1.12	0.405	0.89	0.80 – 0.98	0.014

Appendix 3.8 Continued

				Unadjusted			Adjusted		
	Number (%) Utilising Trastuzumab	Number (%) Not Utilising Trastuzumab	P Value[1]	OR	95% CI	P Value[2]	OR	95% CI	P Value[2]
Women Aged <60 (n = 19,548) Sensitivity Analysis 3									
IMD[3]	n = 11,064 (56.60)	n = 8,484 (43.40)	0.001			**0.0010**			**0.5992**
1 (Least Deprived)	2,407 (54.87)	1,980 (45.13)		1.00	----- -----	-------	1.00	----- -----	-------
2	2,373 (56.07)	1,859 (43.93)		1.05	0.95 – 1.14	0.260	1.04	0.95 – 1.14	0.428
3	2,158 (55.62)	1,722 (44.38)		1.03	0.95 – 1.12	0.493	1.00	0.91 – 1.09	0.936
4	2,159 (58.18)	1,552 (41.82)		1.14	1.05 – 1.25	0.003	1.06	0.96 – 1.17	0.252
5 (Most Deprived)	1,967 (58.93)	1,371 (41.07)		1.18	1.08 – 1.29	<0.001	1.06	0.96 – 1.18	0.269
Women Aged ≥60 (n = 20,631) Sensitivity Analysis 3									
IMD[3]	n = 6,610 (32.04)	n = 14,021 (67.96)	0.000			**<0.001**			**<0.001**
1 (Least Deprived)	1,681 (34.41)	3,204 (65.59)		1.00	----- -----	-------	1.00	----- -----	-------
2	1,586 (33.50)	3,149 (66.50)		0.96	0.88 – 1.04	0.343	0.95	0.86 – 1.04	0.283
3	1,298 (30.56)	2,949 (69.44)		0.84	0.77 – 0.92	<0.001	0.84	0.76 – 0.93	0.001
4	1,107 (30.21)	2,557 (69.79)		0.83	0.75 – 0.90	<0.001	0.79	0.71 – 0.88	<0.001
5 (Most Deprived)	938 (30.26)	2,162 (69.74)		0.83	0.75 – 0.91	<0.001	0.76	0.68 – 0.86	<0.001

[1]Chi-square P value
[2]P Values in bold are from likelihood ratio tests of the variable's contribution to the model. Unbolded P values are from a test of whether the OR is different from 1.
[3]For diagnosis year 2012, IMD_2010 was used and for diagnosis years 2013 - 2017, IMD_2015 was used
All models are adjusted for: age, diagnosis year, ethnicity, rural/urban categorisation, government region, stage, grade, ER status, comorbidities, and whether discussed at MDT.
HER2: Human epidermal growth factor receptor 2; IMD: Index of multiple deprivation; OR: Odds ratio; SACT: Systemic Anti-Cancer Therapy; 95% CI: 95% Confidence interval.

Appendix 3.9 Trastuzumab multivariable model for patients with only a SACT record for Chapter 3 (Sensitivity Analysis 4)

Likelihood (OR with 95% CI and p values from logistic regression) of receiving trastuzumab by deprivation and adjusted for: age, year of diagnosis, ethnicity, rural/urban categorisation, government region, stage, grade, ER status, comorbidities, and whether discussed at MDT or not for women with HER2+ breast cancer diagnosed between 01/01/2012 – 31/12/2017 (n = 21,881)

				Unadjusted			**Adjusted**		
	Number (%) Receiving Trastuzumab (n = 17,674)	Number (%) Not Receiving Trastuzumab (n = 4,207)	Chi2 P Value	OR	95% CI	P Value	OR	95% CI	P Value
IMD[1]			0.001			**0.001**			**0.006**
1 (Least Deprived)	4,088 (82.42)	872 (17.58)		1.00	----- -----	-------	1.00	----- -----	-------
2	3,959 (81.65)	890 (18.35)		0.95	0.86 – 1.05	0.318	0.97	0.87 – 1.07	0.511
3	3,456 (79.72)	879 (20.28)		0.84	0.76 – 0.93	0.001	0.84	0.75 – 0.93	0.001
4	3,266 (79.93)	820 (20.07)		0.85	0.76 – 0.94	0.003	0.86	0.77 – 0.97	0.010
5 (Most Deprived)	2,905 (79.57)	746 (20.43)		0.83	0.75 – 0.93	0.001	0.87	0.78 – 0.98	0.025
Age (Years)			<0.001			**<0.001**			**<0.001**
<40	1,707 (84.55)	312 (15.45)		4.50	3.64 – 5.57	<0.001	3.85	3.08 – 4.82	<0.001
40 – 49	4,158 (83.13)	844 (16.87)		4.06	3.35 – 4.90	<0.001	3.48	2.84 – 4.26	<0.001
50 – 59	5,199 (82.04)	1,138 (17.96)		3.76	3.12 – 4.53	<0.001	3.24	2.66 – 3.95	<0.001
60 – 69	4,372 (80.29)	1,073 (19.71)		3.35	2.78 – 4.05	<0.001	2.81	2.31 – 3.42	<0.001
70 – 79	1,961 (76.21)	612 (23.79)		2.64	2.17 – 3.21	<0.001	2.33	1.89 – 2.86	<0.001
80+	277 (54.85)	228 (45.15)		1.00	----- -----	-------	1.00	----- -----	-------
Diagnosis Year			<0.001			**<0.001**			**<0.001**
2012	1,567 (85.16)	273 (14.84)		1.83	1.59 – 2.12	<0.001	1.76	1.52 – 2.05	<0.001
2013	2,575 (85.12)	450 (14.88)		1.83	1.62 – 2.06	<0.001	1.72	1.51 – 1.94	<0.001
2014	2,866 (83.46)	568 (16.54)		1.61	1.44 – 1.80	<0.001	1.53	1.36 – 1.72	<0.001
2015	3,235 (81.63)	728 (18.37)		1.42	1.28 – 1.57	<0.001	1.35	1.21 – 1.50	<0.001
2016	3,651 (78.82)	981 (21.18)		1.19	1.08 – 1.31	<0.001	1.16	1.05 – 1.28	0.003
2017	3,780 (75.80)	1,207 (24.20)		1.00	----- -----	-------	1.00	----- -----	-------

Appendix 3.9 Continued

				Unadjusted			Adjusted		
	Number (%) Receiving Trastuzumab (n = 17,674)	Number (%) Not Receiving Trastuzumab (n = 4,207)	Chi2 P Value	OR	95% CI	P Value	OR	95% CI	P Value
Ethnicity			0.025			**0.025**			0.009
White	15,566 (80.68)	3,728 (19.32)		1.00	----- -----	-------	1.00	----- -----	-------
Other Ethnic Group[2]	1,602 (82.62)	337 (17.38)		1.14	1.01 – 1.29	0.038	1.03	0.90 – 1.18	0.647
Missing/unknown[3]	506 (78.09)	142 (21.91)		0.85	0.71 – 1.03	1.101	0.73	0.60 – 0.89	0.002
Rural/Urban Indicator			0.062			0.0626			0.0002
Rural Village, Hamlet & Isolated Dwellings	1,923 (81.28)	443 (18.72)		0.98	0.87 – 1.11	0.766	0.83	0.72 – 0.96	0.010
Rural Town & Fringe	1,856 (79.32)	484 (20.68)		0.87	0.77 – 0.97	0.016	0.78	0.68 – 0.89	<0.001
Urban City & Town	7,787 (80.40)	1,898 (19.60)		0.93	0.86 – 1.00	0.058	0.81	0.74 – 0.90	<0.001
Urban Conurbation	6,108 (81.55)	1,382 (18.45)		1.00	----- -----	-------	1.00	----- -----	-------
Government Region			<0.001			**<0.001**			<0.001
North West	2,795 (75.72)	896 (24.28)		1.00	----- -----	-------	1.00	----- -----	-------
North East	981 (70.73)	406 (29.27)		0.77	0.67 – 0.89	<0.001	0.85	0.73 – 0.98	0.023
West Midlands	1,725 (81.21)	399 (18.79)		1.39	1.21 – 1.58	<0.001	1.50	1.30 – 1.72	<0.001
Yorkshire & the Humber	1,943 (91.26)	186 (8.74)		3.35	2.83 – 3.96	<0.001	3.26	2.75 – 3.88	<0.001
East Midlands	1,699 (75.01)	566 (24.99)		0.96	0.85 – 1.09	0.534	1.01	0.89 – 1.15	0.858
East of England	2,161 (79.77)	548 (20.23)		1.26	1.12 – 1.43	<0.001	1.43	1.25 – 1.63	<0.001
South East	2,773 (85.09)	486 (14.91)		1.83	1.62 – 2.07	<0.001	2.03	1.77 – 2.32	<0.001
South West	1,881 (84.54)	344 (15.46)		1.75	1.53 – 2.01	<0.001	2.06	1.77 – 2.40	<0.001
London	1,716 (82.03)	376 (17.97)		1.46	1.28 – 1.67	<0.001	1.34	1.15 – 1.56	<0.001
Stage			0.001			**0.001**			<0.001
I	4,680 (82.66)	982 (17.34)		1.00	----- -----	-------	1.00	----- -----	-------
II	8,189 (80.38)	1,999 (19.62)		0.86	0.79 – 0.94	<0.001	0.83	0.77 – 0.91	<0.001
III	2,698 (79.47)	697 (20.53)		0.81	0.73 – 0.90	<0.001	0.79	0.71 – 0.88	<0.001
IV	1,062 (79.91)	267 (20.09)		0.83	0.72 – 0.97	0.019	0.85	0.72 – 1.00	0.044
Unknown[4]	1,045 (79.95)	262 (20.05)		0.84	0.72 – 0.97	0.022	0.76	0.65 – 0.89	0.001

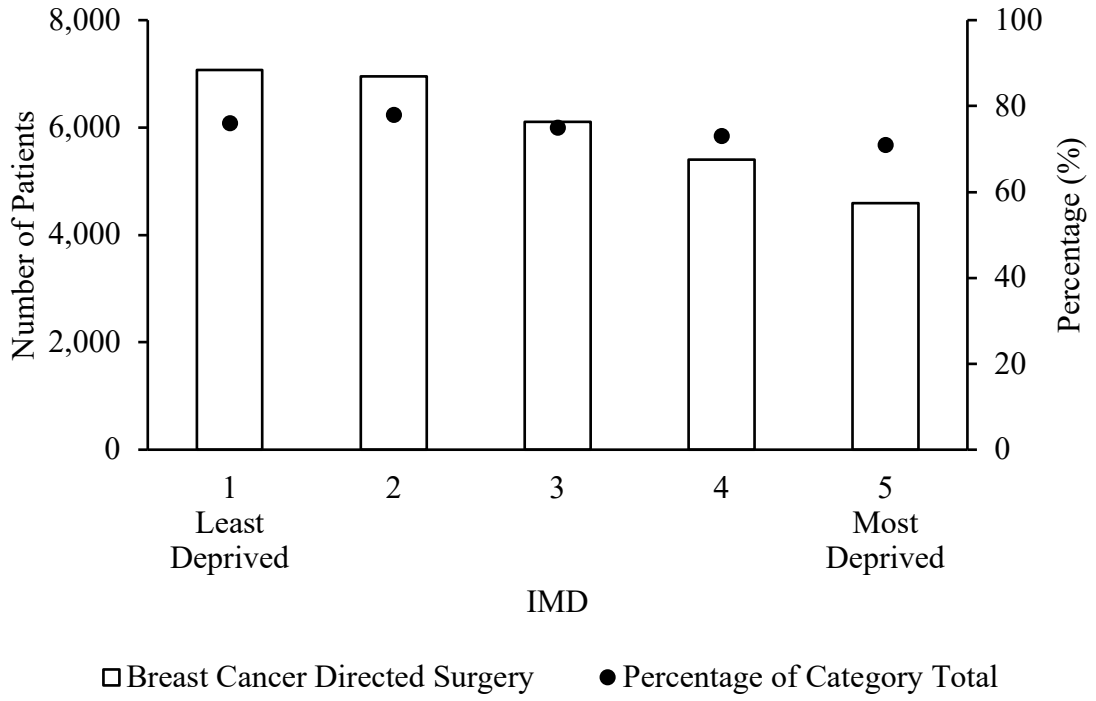
Appendix 3.9 Continued

				Unadjusted			Adjusted		
	Number (%) Receiving Trastuzumab (n = 17,674)	Number (%) Not Receiving Trastuzumab (n = 4,207)	Chi2 P Value	OR	95% CI	P Value	OR	95% CI	P Value
Grade			<0.001			**<0.001**			-------
Well Differentiated (Low Grade)	251 (58.78)	176 (41.22)		0.27	0.22 – 0.33	<0.001	-----	----- -----	-------
Moderately Differentiated	6,302 (76.68)	1,917 (23.32)		0.63	0.59 – 0.67	<0.001	-----	----- -----	-------
Poorly Differentiated	10,770 (83.96)	2,057 (16.04)		1.00	----- -----	-------	-----	----- -----	-------
Other[5]	351 (86.03)	57 (13.97)		1.18	0.89 – 1.56	0.263	-----	----- -----	-------
Big Tumour Count			<0.001			**<0.001**			-------
1	15,999 (81.52)	3,626 (18.48)		1.00	----- -----	-------	-----	----- -----	-------
>1	1,675 (74.25)	581 (25.75)		0.65	0.59 – 0.72	<0.001	-----	----- -----	-------
ER Status			<0.001			**<0.001**			**<0.001**
Positive	9,878 (78.08)	2,773 (21.92)		1.00	----- -----	-------	1.00	----- -----	-------
Negative	4,855 (86.91)	731 (13.09)		1.86	1.71 – 2.04	<0.001	1.91	1.74 – 2.08	<0.001
Unknown	2,941 (80.71)	703 (19.29)		1.17	1.07 – 1.29	0.001	1.19	1.08 – 1.32	0.001
CCM (78 to 6 Months Prior to Diagnosis)			<0.001			**<0.001**			**0.001**
0	15,414 (81.46)	3,508 (18.54)		1.00	----- -----	-------	1.00	----- -----	-------
1-2	2,044 (77.31)	600 (22.69)		0.76	0.70 – 0.86	<0.001	0.85	0.79 – 0.97	0.013
3+	216 (68.57)	99 (31.43)		0.50	0.39 – 0.63	<0.001	1.18	0.50 – 0.84	0.001
Discussed at MDT			<0.001			**<0.001**			**<0.001**
Yes	12,936 (81.49)	2,939 (18.51)		1.00	----- -----	-------	1.00	----- -----	-------
No	2,536 (75.93)	804 (24.07)		0.72	0.66 – 0.78	<0.001	0.85	0.77 – 0.93	0.001
Missing	2,202 (82.60)	464 (17.40)		1.08	0.97 – 1.20	0.171	1.18	1.05 – 1.32	0.005

[1]For diagnosis year 2012, IMD_2010 was used and for diagnosis years 2013 - 2017, IMD_2015 was used.
[2]Other refers to Asian/British Asian, Black/African/Caribbean/Black British, mixed/multiple ethnic groups and other ethnic groups.
[3]Missing/unknown refers to missing and unknown ethnicity classifications.
[4]Unknown staging refers to missing and unstageable tumours
[5]Other refers to undifferentiated or anaplastic, undetermined, and missing tumour grades. CCM, Charlson comorbidity index; ER, Oestrogen receptor status; IMD: Index of multiple deprivation. MDT: Multi-disciplinary team; OR, Odds ratio; 95% CI, 95% Confidence interval.

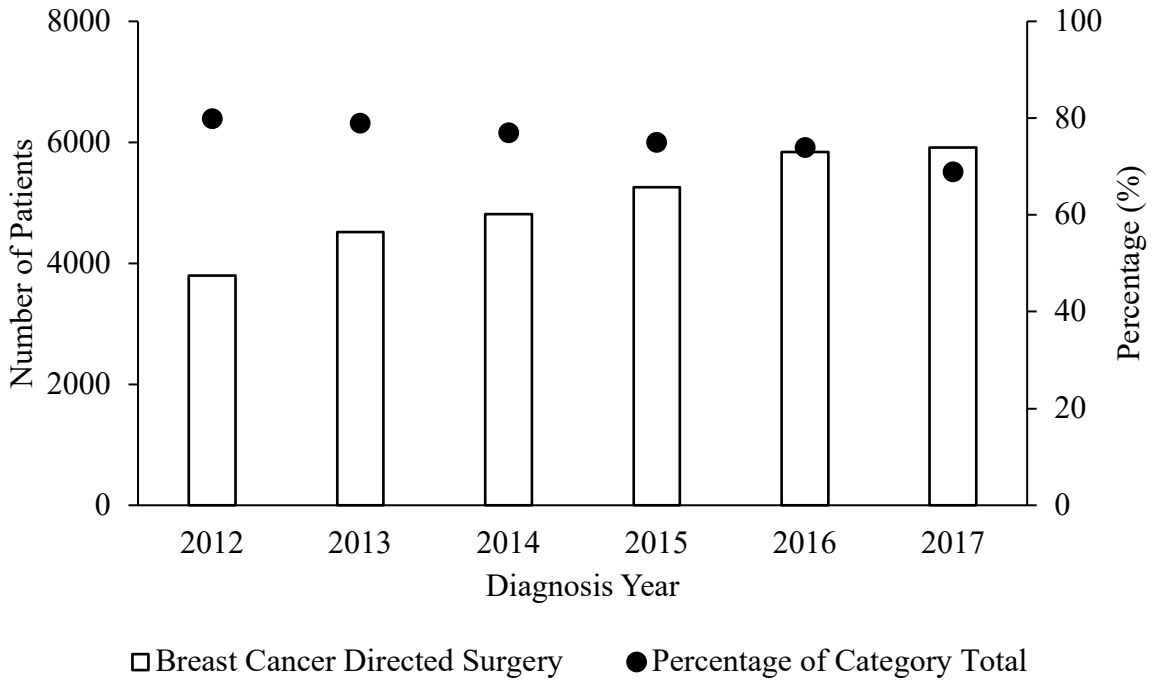
Appendix 3.10 Breast cancer directed surgery utilisation by a) deprivation; b) year of diagnosis and c) stage for Chapter 3.

a)



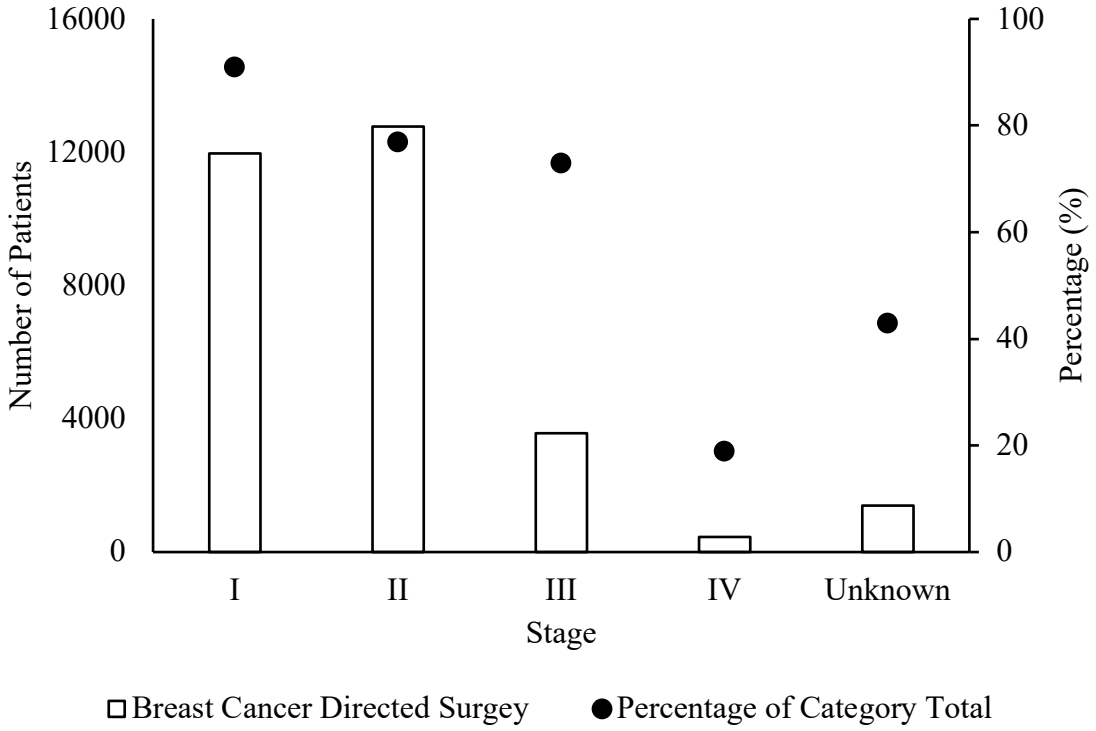
IMD: Index of multiple deprivation

b)



Appendix 3.10 Continued

c)



Appendix 3.11 Sensitivity analyses for breast cancer directed surgery utilisation for Chapter 3

				Unadjusted			Adjusted		
	Number (%) Utilising Surgery	Number (%) Not Utilising Surgery	P Value[1]	OR	95% CI	P Value[2]	OR	95% CI	P Value[2]
Full Analytical Cohort Original Model for Comparison (n = 40,179)									
IMD[3]	n = 30,124 (74.97)	n = 10,055 (25.03)	<0.001			**<0.001**			**<0.001**
1 (Least Deprived)	7,075 (76.31)	2,197 (23.69)		1.00	----- -----	-------	1.00	----- -----	-------
2	6,948 (77.48)	2,019 (22.52)		1.07	1.00 – 1.14	0.059	1.02	0.94 – 1.10	0.618
3	6,109 (75.17)	2,018 (24.83)		0.94	0.88 – 1.01	0.081	0.91	0.84 – 0.98	0.016
4	5,397 (73.18)	1,978 (26.82)		0.85	0.79 – 0.91	<0.001	0.86	0.79 – 0.93	<0.001
5 (Most Deprived)	4,595 (71.37)	1,843 (28.63)		0.77	0.72 – 0.83	<0.001	0.79	0.73 – 0.86	<0.001
Diagnosis Date Post Mandated SACT Submission 01/04/2014 – 31/12/2017 (n = 28,146) Sensitivity Analysis 1									
IMD[3]	n = 20,594 (73.17)	n = 7,552 (26.83)	<0.001			**<0.001**			**<0.001**
1 (Least Deprived)	4,820 (74.31)	1,666 (25.69)		1.00	----- -----	-------	1.00	----- -----	-------
2	4,776 (75.77)	1,527 (24.23)		1.08	1.00 – 1.17	0.057	1.03	0.94 – 1.13	0.473
3	4,168 (73.57)	1,497 (26.43)		0.96	0.89 – 1.04	0.354	0.94	0.86 – 1.03	0.208
4	3,700 (71.06)	1,507 (28.94)		0.85	0.78 – 0.92	<0.001	0.85	0.78 – 0.94	0.001
5 (Most Deprived)	3,130 (69.79)	1,355 (30.21)		0.80	0.73 – 0.87	<0.001	0.82	0.74 – 0.90	<0.001
Positive HER2+ Status Definition (n = 27,712) Sensitivity Analysis 2									
IMD[3]	n = 20,666 (74.57)	n = 7,046 (25.43)	<0.001			**<0.001**			**<0.001**
1 (Least Deprived)	4,876 (75.82)	1,555 (24.18)		1.00	----- -----	-------	1.00	----- -----	-------
2	4,798 (77.20)	1,417 (22.80)		1.08	0.99 – 1.17	0.067	1.04	0.95 – 1.14	0.430
3	4,157 (74.67)	1,410 (25.33)		0.94	0.87 – 1.02	0.146	0.92	0.83 – 1.01	0.069
4	3,723 (73.22)	1,362 (26.78)		0.87	0.80 – 0.95	0.001	0.91	0.83 – 1.00	0.061
5 (Most Deprived)	3,112 (70.50)	1,302 (29.50)		0.76	0.70 – 0.83	<0.001	0.80	0.72 – 0.88	<0.001

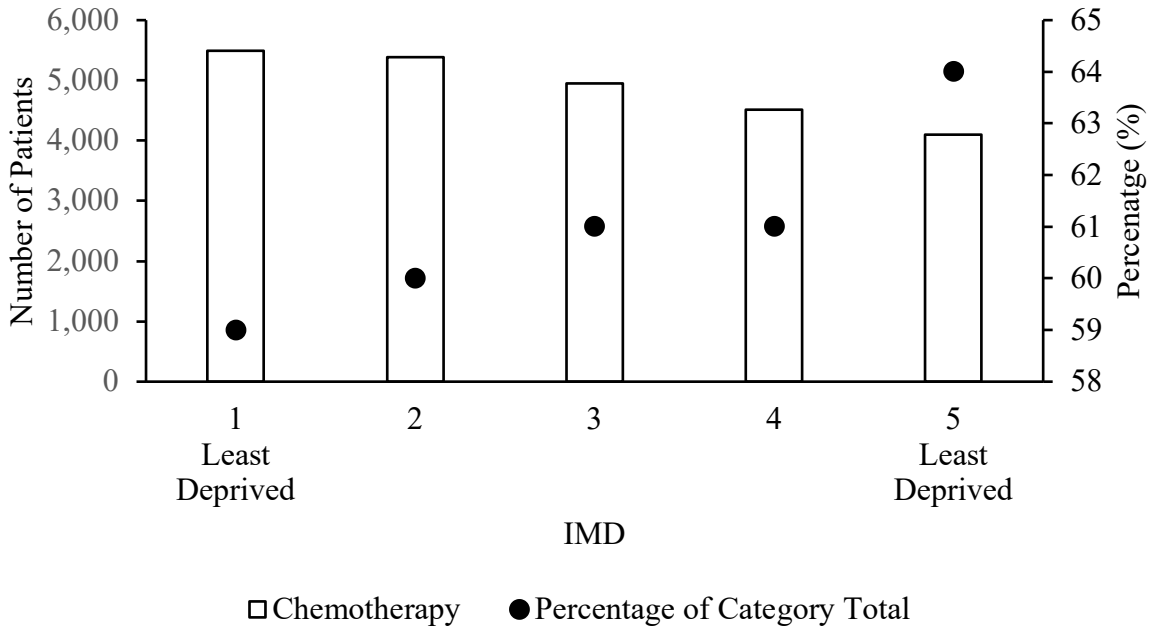
Appendix 3.11 Continued

				Unadjusted			Adjusted		
	Number (%) Utilising Surgery	Number (%) Not Utilising Surgery	P Value[1]	OR	95% CI	P Value[2]	OR	95% CI	P Value[2]
Women Aged <60 (n = 19,548) Sensitivity Analysis 3									
IMD[3]	n = 14,922 (76.34)	n = 4,626 (23.66)	<0.001			**<0.001**			**0.027**
1 (Least Deprived)	3,370 (76.82)	1,017 (23.18)		1.00	----- -----	-------	1.00	----- -----	-------
2	3,349 (79.14)	883 (20.86)		1.14	1.03 – 1.27	0.010	1.06	0.95 – 1.19	0.288
3	2,975 (76.68)	905 (23.32)		0.99	0.90 – 1.10	0.878	0.95	0.85 – 1.07	0.395
4	2,767 (74.56)	944 (25.44)		0.88	0.80 – 0.98	0.018	0.91	0.81 – 1.02	0.120
5 (Most Deprived)	2,461 (73.73)	877 (26.27)		0.85	0.76 – 0.94	0.002	0.88	0.78 – 1.00	0.042
Women Aged ≥60 (n = 20,631) Sensitivity Analysis 3									
IMD[3]	n = 15,202 (73.69)	n = 5,429 (26.31)	<0.001			**<0.001**			**<0.001**
1 (Least Deprived)	3,705 (75.84)	1,180 (24.16)		1.00	----- -----	-------	1.00	----- -----	-------
2	3,599 (76.01)	1,136 (23.99)		1.01	0.92 – 1.11	0.851	0.98	0.88 – 1.09	0.674
3	3,134 (73.79)	1,113 (26.21)		0.90	0.82 – 0.99	0.024	0.87	0.78 – 0.97	0.011
4	2,630 (71.78)	1,034 (28.22)		0.81	0.73 – 0.89	<0.001	0.82	0.73 – 0.92	0.001
5 (Most Deprived)	2,134 (68.84)	966 (31.16)		0.70	0.64 – 0.78	<0.001	0.71	0.63 – 0.80	<0.001

[1]Chi-square P value
[2]P Values in bold are from likelihood ratio tests of the variable's contribution to the model. Unbolded P values are from a test of whether the OR is different from 1.
[3]For diagnosis year 2012, IMD_2010 was used and for diagnosis years 2013 - 2017, IMD_2015 was used
All models adjusted for: age and stage, diagnosis year, ethnicity, rural/urban categorisation, grade, ER status, comorbidities, and whether discussed at MDT.
HER2: Human epidermal growth factor receptor 2; IMD: Index of multiple deprivation; OR: Odds ratio; SACT: Systemic Anti-Cancer Therapy; 95% CI: 95% Confidence interval.

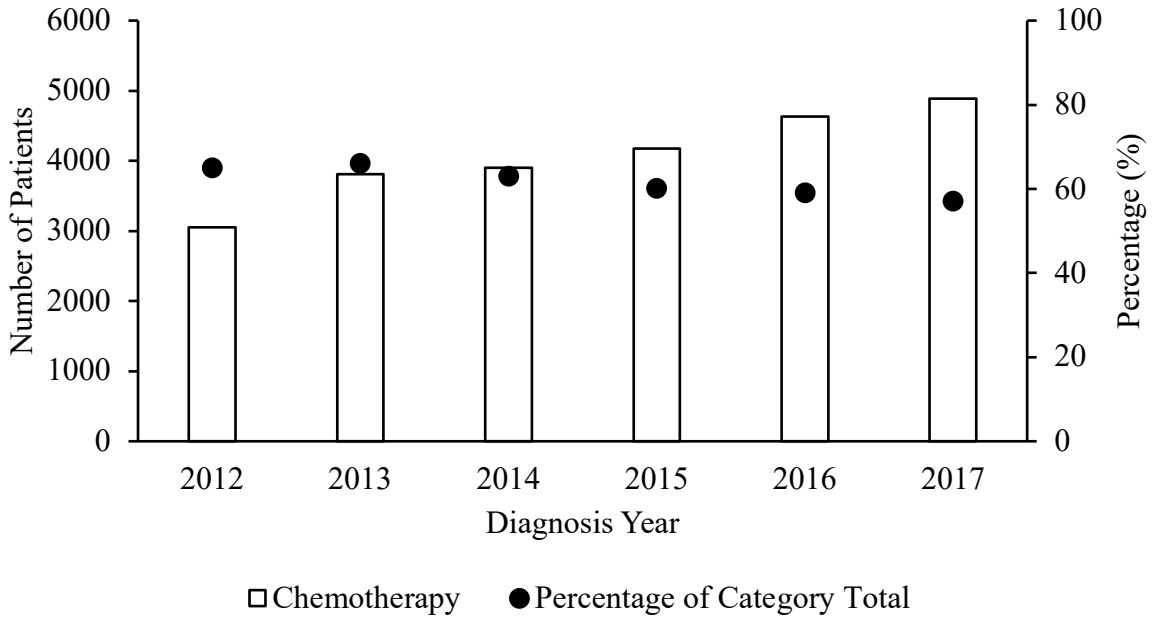
Appendix 3.12 Chemotherapy utilisation by a) deprivation; b) year of diagnosis and c) stage for Chapter 3.

a)



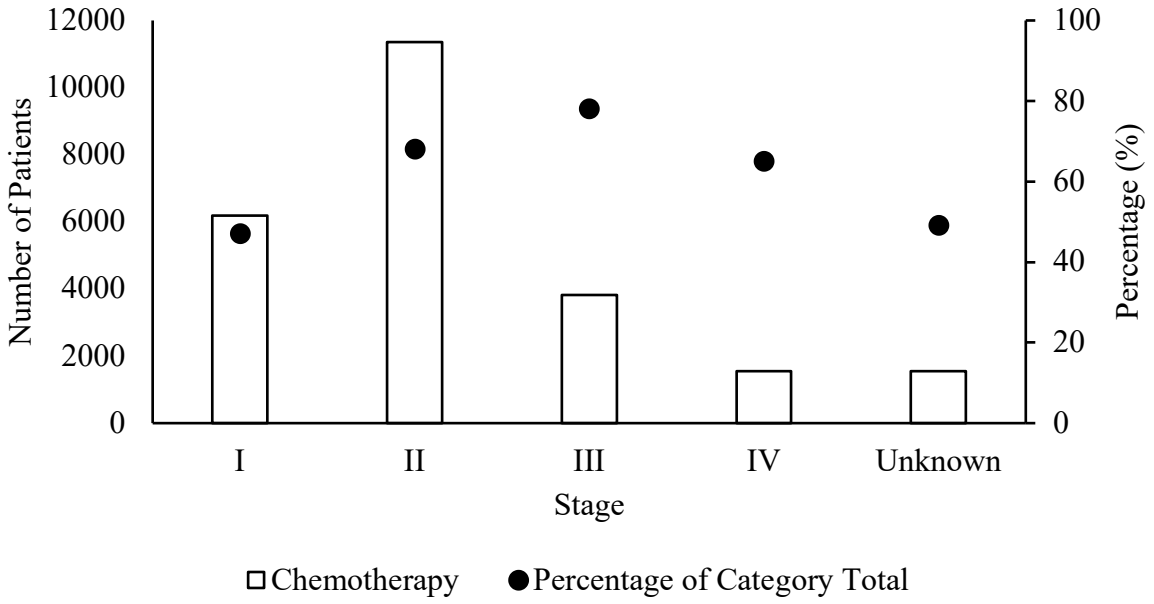
IMD: Index of multiple deprivation.

b)



Appendix 3.12 Continued

c)



Appendix 3.13 Chemotherapy models (older women, had surgery) for Chapter 3

Likelihood (OR with 95% CI and p values from logistic regression) of utilising chemotherapy in women by deprivation and adjusted for diagnosis year, stage, ethnicity, comorbidities and whether discussed at MDT if receiving surgery and aged ≥60 with HER2+ breast cancer diagnosed between 01/01/2012 – 31/12/2017 (n = 15,202)

				Unadjusted			**Adjusted**		
	Number (%) Utilising Chemotherapy n = 7,236 (47.60)	Number (%) Not Utilising Chemotherapy n = 7,966 (56.40)	P Value[1]	OR	95% CI	P Value[2]	OR	95% CI	P Value[2]
IMD[3]			0.715			**0.715**			**0.718**
1 (Least Deprived)	1,777 (47.96)	1,928 (52.04)		1.00	----- -----	-------	1.00	----- -----	-------
2	1,732 (48.12)	1,867 (51.88)		1.00	0.92 – 1.10	0.890	1.02	0.93 – 1.13	0.627
3	1,500 (47.86)	1,634 (52.14)		1.00	0.91 – 1.10	0.934	1.02	0.92 – 1.12	0.729
4	1,233 (46.88)	1,397 (53.12)		0.96	0.87 – 1.06	0.396	0.98	0.88 – 1.09	0.675
5 (Most Deprived)	994 (46.58)	1,140 (53.42)		0.95	0.85 – 1.05	0.308	0.95	0.85 – 1.07	0.396
Diagnosis Year			<0.001			**<0.001**			**<0.001**
2012	914 (50.67)	890 (49.33)		1.42	1.26 – 1.59	<0.001	1.26	1.12 – 1.43	<0.001
2013	1,150 (52.56)	1,038 (47.44)		1.53	1.37 – 1.71	<0.001	1.37	1.22 – 1.54	<0.001
2014	1,233 (50.45)	1,211 (49.55)		1.41	1.26 – 1.56	<0.001	1.28	1.14 – 1.43	<0.001
2015	1,278 (47.35)	1,421 (52.65)		1.24	1.12 – 1.38	<0.001	1.17	1.05 – 1.30	0.006
2016	1,370 (45.77)	1,623 (54.23)		1.17	1.05 – 1.29	0.003	1.16	1.04 – 1.28	0.007
2017	1,291 (42.00)	1,783 (58.00)		1.00	----- -----	-------	1.00	----- -----	-------
Ethnicity			<0.001			**<0.001**			**<0.001**
White	6,655 (47.75)	7,282 (52.25)		1.00	----- -----	-------	1.00	----- -----	-------
Other Ethnic Group[4]	365 (51.99)	337 (48.01)		1.19	1.02 – 1.38	0.028	1.26	1.08 – 1.48	0.004
Missing/unknown[5]	216 (38.37)	347 (61.63)		0.68	0.57 – 0.81	<0.001	0.65	0.54 – 0.78	<0.001
Rural/Urban Indicator			0.004			**0.004**			-------
Rural Village, Hamlet & Isolated Dwellings	886 (49.69)	897 (50.31)		1.07	0.96 – 1.19	0.232	-----	----- -----	-------
Rural Town & Fringe	888 (50.03)	887 (49.97)		1.08	0.97 – 1.21	0.151	-----	----- -----	-------
Urban City & Town	3,249 (46.18)	3,787 (53.82)		0.93	0.86 – 1.00	0.051	-----	----- -----	-------
Urban Conurbation	2,213 (48.03)	2,395 (51.97)		1.00	----- -----	-------	-----	----- -----	-------

Appendix 3.13 Continued

				Unadjusted			Adjusted		
	Number (%) Utilising Chemotherapy n = 7,236 (47.60)	Number (%) Not Utilising Chemotherapy n = 7,966 (56.40)	P Value[1]	OR	95% CI	P Value[2]	OR	95% CI	P Value[2]
Government Region			<0.001			**<0.001**			-------
North West	1,075 (43.88)	1,375 (56.12)		1.00	----- -----	-------	-----	----- -----	-------
North East	514 (51.40)	486 (48.60)		1.35	1.17 – 1.57	<0.001	-----	----- -----	-------
West Midlands	832 (50.58)	813 (49.42)		1.31	1.15 – 1.48	<0.001	-----	----- -----	-------
Yorkshire & the Humber	720 (54.92)	591 (45.08)		1.56	1.36 – 1.78	<0.001	-----	----- -----	-------
East Midlands	650 (46.03)	762 (53.97)		1.09	0.96 – 1.24	0.194	-----	----- -----	-------
East of England	954 (48.87)	998 (51.13)		1.22	1.09 – 1.38	0.001	-----	----- -----	-------
South East	1,124 (45.43)	1,350 (54.57)		1.06	0.95 – 1.19	0.272	-----	----- -----	-------
South West	868 (47.64)	954 (52.36)		1.16	1.03 – 1.31	0.015	-----	----- -----	-------
London	499 (43.93)	637 (56.07)		1.00	0.87 – 1.15	0.978	-----	----- -----	-------
Stage			<0.001			**<0.001**			**<0.001**
I	2,283 (36.41)	3,987 (63.59)		1.00	----- -----	-------	1.00	----- -----	-------
II	3,342 (52.79)	2.989 (47.21)		1.95	1.82 – 2.10	<0.001	1.99	1.85 – 2.13	<0.001
III	1,154 (65.98)	595 (34.02)		3.39	3.03 – 3.79	<0.001	3.56	3.18 – 3.99	<0.001
IV	133 (61.29)	84 (38.71)		2.77	2.09 – 3.65	<0.001	2.85	2.15 – 3.78	<0.001
Unknown[6]	324 (51.02)	311 (48.98)		1.82	1.54 – 2.14	<0.001	1.82	1.53 – 2.15	<0.001
Grade			<0.001			**<0.001**			-------
Well Differentiated (Low Grade)	123 (12.89)	831 (87.11)		0.09	0.08 – 0.11	<0.001	-----	----- -----	-------
Moderately Differentiated	2,575 (37.96)	4,209 (62.04)		0.39	0.36 – 0.42	<0.001	-----	----- -----	-------
Poorly Differentiated	4,464 (61.06)	2,847 (38.94)		1.00	----- -----	-------	-----	----- -----	-------
Other[7]	74 (48.37)	79 (51.63)		0.60	0.43 – 0.82	0.002	-----	----- -----	-------
Big Tumour Count			<0.001			**<0.001**			-------
1	6,197 (48.83)	6,494 (51.17)		0.74	0.68 – 0.81	<0.001	-----	----- -----	-------
>1	1,039 (41.38)	1,472 (58.62)		0.95	0.92 – 0.99	0.008	-----	----- -----	-------

Appendix 3.13 Continued

				Unadjusted			Adjusted		
	Number (%) Utilising Chemotherapy n = 7,236 (47.60)	Number (%) Not Utilising Chemotherapy n = 7,966 (56.40)	P Value[1]	OR	95% CI	P Value[2]	OR	95% CI	P Value[2]
ER Status			<0.001			**<0.001**			-------
Positive	4,120 (42.33)	5,612 (57.67)		1.00	----- -----	-------	-----	----- -----	-------
Negative	2,123 (62.20)	1,290 (37.80)		2.24	2.07 – 2.43	<0.001	-----	----- -----	-------
Unknown	993 (48.27)	1,064 (51.73)		1.27	1.16 – 1.40	<0.001	-----	----- -----	-------
CCM (Between 78 to 6 Months Prior to Diagnosis)			<0.001			**<0.001**			**<0.001**
0	5,922 (50.84)	5,727 (49.16)		1.00	----- -----	-------	1.00	----- -----	-------
1-2	1,179 (40.32)	1,745 (59.68)		0.65	0.60 – 0.71	<0.001	0.62	0.57 – 0.67	<0.001
3+	135 (21.46)	494 (78.54)		0.26	0.22 – 0.32	<0.001	0.24	0.20 – 0.29	<0.001
Discussed at MDT			<0.001			**<0.001**			**<0.001**
Yes	5,616 (49.64)	5,697 (50.36)		1.00	----- -----	-------	1.00	----- -----	-------
No	972 (41.33)	1,380 (58.67)		0.71	0.65 – 0.78	<0.001	0.78	0.71 – 0.85	<0.001
Missing	648 (42.16)	889 (57.84)		0.74	0.66 – 0.82	<0.001	0.72	0.64 – 0.80	<0.001

[1]Chi-square P value
[2]P Values in bold are from likelihood ratio tests of the variables' contribution to the model. Unbolded P values are from a test of whether the OR is different from 1.
[3]For diagnosis year 2010, IMD_2010 was used and for diagnosis years 2013 - 2017, IMD_2015 was used.
[4]Other refers to Asian/British Asian, Black/African/Caribbean/Black British, mixed/multiple ethnic groups and other ethnic groups.
[5]Missing/unknown refers to missing and unknown ethnicity classifications.
[6]Unknown staging refers to missing and unstageable tumours.
[7]Other refers to undifferentiated or anaplastic, undetermined, and missing tumour grades.
CCM: Charlson Comorbidity Index; ER: Oestrogen receptor status; IMD: Index of multiple deprivation; MDT: Multidisciplinary team; OR: Odds ratio; 95% CI: 95% Confidence interval.

Appendix 3.14 Chemotherapy models (younger women, had surgery) for Chapter 3

Likelihood (OR with 95% CI and p values from logistic regression) of utilising chemotherapy in women by deprivation and adjusted for: diagnosis year, stage, ethnicity, and whether discussed at MDT if receiving surgery and aged <60 with a HER2+ breast cancer diagnosed between 01/01/2012 – 31/12/2017 (n = 14,922)

				Unadjusted			**Adjusted**		
	Number (%) Utilising Chemotherapy n = 11,518 (77.19)	Number (%) Not Utilising Chemotherapy n = 3,404 (22.81)	P Value[1]	OR	95% CI	P Value[2]	OR	95% CI	P Value[2]
IMD[3]			<0.001			**<0.001**			**0.001**
1 (Least Deprived)	2,488 (73.83)	882 (26.17)		1.00	----- -----	-------	1.00	----- -----	-------
2	2,564 (76.56)	785 (23.44)		1.16	1.04 – 1.29	0.010	1.13	1.00 – 1.27	0.054
3	2,299 (77.28)	676 (22.72)		1.21	1.07 – 1.35	0.001	1.12	0.99 – 1.27	0.071
4	2,163 (78.17)	604 (21.83)		1.27	1.13 – 1.43	<0.001	1.14	1.00 – 1.29	0.054
5 (Most Deprived)	2,004 (81.43)	457 (18.57)		1.55	1.37 – 1.77	<0.001	1.36	1.19 – 1.57	<0.001
Diagnosis Year			<0.001			**<0.001**			**<0.001**
2012	1,623 (81.60)	366 (18.40)		2.04	1.78 – 2.35	<0.001	1.93	1.66 – 2.24	<0.001
2013	1,961 (84.27)	366 (15.73)		2.47	2.15 – 2.83	<0.001	2.40	2.07 – 2.78	<0.001
2014	1,901 (80.31)	466 (19.69)		1.88	1.65 – 2.14	<0.001	1.81	1.57 – 2.08	<0.001
2015	2,004 (78.28)	556 (21.72)		1.66	1.47 – 1.88	<0.001	1.63	1.42 – 1.86	<0.001
2016	2,088 (73.44)	755 (26.56)		1.28	1.14 – 1.43	<0.001	1.33	1.17 – 1.51	<0.001
2017	1,941 (68.44)	895 (31.56)		1.00	----- -----	-------	1.00	----- -----	-------
Ethnicity			<0.001			**<0.001**			**<0.001**
White	10,098 (78.48)	2,769 (21.52)		1.00	----- -----	-------	1.00	----- -----	-------
Other Ethnic Group[4]	1,069 (76.09)	336 (23.91)		0.87	0.77 – 0.99	0.039	0.80	0.69 – 0.92	0.002
Missing/unknown[5]	351 (54.00)	299 (46.00)		0.32	0.27 – 0.38	<0.001	0.39	0.32 – 0.46	<0.001
Rural/Urban Indicator			0.028			**0.031**			-------
Rural Village, Hamlet & Isolated Dwellings	1,233 (74.28)	427 (25.72)		0.83	0.73 – 0.94	0.004	-----	----- -----	-------
Rural Town & Fringe	1,170 (77.33)	343 (22.67)		0.98	0.85 – 1.12	0.753	-----	----- -----	-------
Urban City & Town	5,317 (77.48)	1,545 (22.52)		0.99	0.90 – 1.08	0.767	-----	----- -----	-------
Urban Conurbation	3,798 (77.72)	1,089 (22.28)		1.00	----- -----	-------	-----	----- -----	-------

Appendix 3.14 Continued

				Unadjusted			Adjusted		
	Number (%) Utilising Chemotherapy n = 11,518 (77.19)	Number (%) Not Utilising Chemotherapy n = 3,404 (22.81)	P Value[1]	OR	95% CI	P Value[2]	OR	95% CI	P Value[2]
Government Region			<0.001			**<0.001**			-------
North West	1,864 (76.42)	575 (23.58)		1.00	----- -----	-------	-----	----- -----	-------
North East	806 (83.26)	162 (16.74)		1.53	1.27 – 1.86	<0.001	-----	----- -----	-------
West Midlands	1,207 (76.49)	371 (23.51)		1.00	0.86 – 1.17	0.962	-----	----- -----	-------
Yorkshire & the Humber	1,229 (86.73)	188 (13.27)		2.02	1.68 – 2.41	<0.001	-----	----- -----	-------
East Midlands	1,028 (77.70)	295 (22.30)		1.07	0.92 – 1.26	0.375	-----	----- -----	-------
East of England	1,474 (73.70)	526 (26.30)		0.86	0.75 – 0.99	0.037	-----	----- -----	-------
South East	1,763 (75.25)	580 (24.75)		0.94	0.82 – 1.07	0.341	-----	----- -----	-------
South West	1,228 (76.75)	372 (23.25)		1.02	0.88 – 1.18	0.811	-----	----- -----	-------
London	919 (73.29)	335 (26.71)		0.85	0.72 – 0.99	0.036	-----	----- -----	-------
Stage			<0.001			**<0.001**			**<0.001**
I	3,476 (60.99)	2,233 (39.01)		1.00	----- -----	-------	1.00	----- -----	-------
II	5,533 (86.31)	881 (13.69)		4.03	3.69 – 4.41	<0.001	4.34	3.96 – 4.76	<0.001
III	1,716 (94.75)	95 (5.25)		11.55	9.33 – 14.30	<0.001	11.86	9.55 – 14.72	<0.001
IV	202 (88.99)	25 (11.01)		5.17	3.40 – 7.86	<0.001	5.28	3.44 – 8.11	<0.001
Unknown[6]	571 (76.03)	180 (23.97)		2.03	1.70 – 2.42	<0.001	2.27	1.89 – 2.74	<0.001
Grade			<0.001			**<0.001**			-------
Well Differentiated (Low Grade)	234 (27.82)	607 (72.18)		0.05	0.04 – 0.06	<0.001	-----	----- -----	-------
Moderately Differentiated	4,218 (69.28)	1,870 (30.72)		0.28	0.25 – 0.30	<0.001	-----	----- -----	-------
Poorly Differentiated	6,908 (89.02)	852 (10.98)		1.00	----- -----	-------	-----	----- -----	-------
Other[7]	158 (67.81)	75 (32.19)		0.26	0.20 – 0.35	<0.001	-----	----- -----	-------
Big Tumour Count			0.036			**0.039**			-------
1	10,757 (77.38)	3,144 (22.62)		1.00	----- -----	-------	-----	----- -----	-------
>1	761 (74.53)	260 (25.47)		0.86	0.74 – 0.99	0.036	-----	----- -----	-------

Appendix 3.14 Continued

				Unadjusted			Adjusted		
	Number (%) Utilising Chemotherapy n = 11,518 (77.19)	Number (%) Not Utilising Chemotherapy n = 3,404 (22.81)	P Value[1]	OR	95% CI	P Value[2]	OR	95% CI	P Value[2]
ER Status			<0.001			**<0.001**			-------
Positive	7,235 (73.69)	2,583 (26.31)		1.00	----- -----	-------	-----	----- -----	-------
Negative	2,576 (89.13)	314 (10.87)		2.93	2.58 – 3.32	<0.001	-----	----- -----	-------
Unknown	1,707 (77.10)	507 (22.90)		1.20	1.08 – 1.34	0.001	-----	----- -----	-------
CCM (Between 78 to 6 Months Prior to Diagnosis)			<0.001			**<0.001**			-------
0	10,416 (77.38)	3,045 (22.62)		1.00	----- -----	-------	-----	----- -----	-------
1-2	1,033 (77.49)	300 (22.51)		1.01	0.88 – 1.15	0.924	-----	----- -----	-------
3+	69 (53.19)	59 (46.09)		0.34	0.24 – 0.49	<0.001	-----	----- -----	-------
Discussed at MDT			<0.001			**<0.001**			**<0.001**
Yes	8,763 (80.76)	2,088 (19.24)		1.00	----- -----	-------	1.00	----- -----	-------
No	1,562 (74.20)	543 (25.80)		0.69	0.61 – 0.76	<0.001	0.79	0.70 – 0.89	<0.001
Missing	1,193 (60.68)	773 (39.32)		0.37	0.33 – 0.41	<0.001	0.35	0.32 – 0.40	<0.001

[1]Chi-square P value
[2]Bolded P values are from likelihood ratio tests of the variable contribution to the model. Unbolded P values are from a test of whether the OR is different from 1.
[3]For diagnosis year 2012, IMD_2010 was used and for diagnosis years 2013 - 2017, IMD_2015 was used.
[4]Other refers to Asian/British Asian, Black/African/Caribbean/Black British, mixed/multiple ethnic groups and other ethnic groups.
[5]Missing/unknown refers to missing and unknown ethnicity classifications.
[6]Unknown staging refers to missing and unstageable tumours.
[7]Other refers to undifferentiated or anaplastic, undetermined, and missing tumour grades.
CCM, Charlson Comorbidity Index; ER, Oestrogen receptor status; IMD: Index of multiple deprivation; MDT: Multidisciplinary team; OR, Odds ratio; 95% CI, 95% Confidence interval.

Appendix 3.15 Chemotherapy models (older women, no surgery) for Chapter 3

Likelihood (OR with 95% CI and p values from logistic regression) of utilising chemotherapy in women by deprivation and adjusted for: diagnosis year, stage, ethnicity, comorbidities, and whether discussed at MDT or not for women not receiving surgery who are aged 60 or older with HER2+ breast cancer diagnosed between 01/01/2012 – 31/12/2017 (n = 5,429)

				Unadjusted			**Adjusted**		
	Number (%) Utilising Chemotherapy n = 1,912 (35.22)	Number (%) Not Utilising Chemotherapy n = 3,517 (64.78)	P Value[1]	OR	95% CI	P Value[2]	OR	95% CI	P Value[2]
IMD[3]			0.013			**0.013**			**0.006**
1 (Least Deprived)	460 (38.98)	720 (61.02)		1.00	----- -----	-------	1.00	----- -----	-------
2	403 (35.48)	733 (64.52)		0.86	0.73 – 1.02	0.081	0.81	0.68 – 0.97	0.025
3	371 (33.33)	742 (66.67)		0.78	0.66 – 0.93	0.005	0.77	0.64 – 0.92	0.005
4	335 (32.40)	699 (67.60)		0.75	0.63 – 0.89	0.001	0.71	0.59 – 0.86	<0.001
5 (Most Deprived)	343 (35.51)	623 (64.49)		0.86	0.72 – 1.03	0.098	0.85	0.70 – 1.02	0.087
Diagnosis Year			<0.001			**0.001**			**<0.001**
2012	185 (36.20)	326 (63.80)		0.87	0.71 – 1.08	0.206	0.80	0.64 – 1.01	0.063
2013	206 (33.12)	416 (66.88)		0.76	0.63 – 0.93	0.007	0.72	0.58 – 0.90	0.003
2014	274 (34.90)	511 (65.10)		0.83	0.69 – 0.99	0.038	0.79	0.64 – 0.96	0.017
2015	288 (30.35)	661 (69.65)		0.67	0.56 – 0.80	<0.001	0.62	0.52 – 0.75	<0.001
2016	399 (35.00)	741 (65.00)		0.83	0.71 – 0.97	0.023	0.79	0.66 – 0.94	0.008
2017	560 (39.38)	862 (60.62)		1.00	----- -----	-------	1.00	----- -----	-------
Ethnicity			<0.001			**<0.001**			**<0.001**
White	1,689 (34.79)	3,166 (65.21)		1.00	----- -----	-------	1.00	----- -----	-------
Other Ethnic Group[4]	144 (54.14)	122 (45.86)		2.21	1.73 – 2.84	<0.001	2.31	1.76 – 3.03	<0.001
Missing/unknown[5]	79 (25.65)	229 (74.35)		0.65	0.50 – 0.84	0.001	0.55	0.42 – 0.73	<0.001
Rural/Urban Indicator			<0.001			**<0.001**			
Rural Village, Hamlet & Isolated Dwellings	221 (40.93)	319 (59.07)		1.12	0.92 – 1.36	0.257	-----	----- -----	-------
Rural Town & Fringe	206 (33.66)	406 (66.34)		0.82	0.68 – 0.99	0.043	-----	----- -----	-------
Urban City & Town	784 (32.09)	1,659 (67.91)		0.76	0.67 – 0.87	<0.001	-----	----- -----	-------
Urban Conurbation	701 (38.22)	1,133 (61.78)		1.00	----- -----	-------	-----	----- -----	-------

Appendix 3.15 Continued

				Unadjusted			Adjusted		
	Number (%) Utilising Chemotherapy n = 1,912 (35.22)	Number (%) Not Utilising Chemotherapy n = 3,517 (64.78)	P Value[1]	OR	95% CI	P Value[2]	OR	95% CI	P Value[2]
Government Region			0.003			**0.0037**			-------
North West	281 (32.52)	583 (67.48)		1.00	----- -----	-------	-----	----- -----	-------
North East	112 (31.37)	245 (68.63)		0.95	0.73 – 1.24	0.695	-----	----- -----	-------
West Midlands	205 (34.17)	395 (65.83)		1.08	0.86 – 1.34	0.511	-----	----- -----	-------
Yorkshire & the Humber	154 (36.93)	263 (63.07)		1.21	0.95 – 1.55	0.119	-----	----- -----	-------
East Midlands	179 (35.87)	320 (64.13)		1.16	0.92 – 1.46	0.208	-----	----- -----	-------
East of England	260 (38.52)	415 (61.48)		1.30	1.05 – 1.60	0.015	-----	----- -----	-------
South East	304 (34.35)	581 (65.65)		1.09	0.89 – 1.32	0.418	-----	----- -----	-------
South West	186 (32.01)	395 (67.99)		0.98	0.78 – 1.22	0.839	-----	----- -----	-------
London	231 (41.92)	320 (58.08)		1.50	1.20 – 1.87	<0.001	-----	----- -----	-------
Stage			<0.001			**<0.001**			**<0.001**
I	130 (18.90)	558 (81.10)		1.00	----- -----	-------	1.00	----- -----	-------
II	659 (35.20)	1,213 (64.80)		2.33	1.88 – 2.89	<0.001	2.29	1.84 – 2.86	<0.001
III	329 (52.81)	294 (47.19)		4.80	3.75 – 6.15	<0.001	4.30	3.32 – 5.55	<0.001
IV	565 (49.52)	576 (50.48)		4.21	3.37 – 5.26	<0.001	3.97	3.15 – 5.01	<0.001
Unknown[6]	229 (20.72)	876 (79.28)		1.12	0.88 – 1.43	0.347	1.15	0.89 – 1.48	0.281
Grade			<0.001			**<0.001**			-------
Well Differentiated (Low Grade)	21 (7.64)	254 (92.36)		0.09	0.06 – 0.14	<0.001	-----	----- -----	-------
Moderately Differentiated	769 (28.13)	1,965 (71.87)		0.43	0.38 – 0.48	<0.001	-----	----- -----	-------
Poorly Differentiated	1,025 (47.83)	1,118 (52.17)		1.00	----- -----	-------	-----	----- -----	-------
Other[7]	97 (35.02)	180 (64.98)		0.59	0.45 – 0.76	<0.001	-----	----- -----	-------
Big Tumour Count			0.391			**0.3901**			-------
1	1,567 (35.48)	2,849 (64.52)		1.00	----- -----	-------	-----	----- -----	-------
>1	345 (34.06)	668 (65.94)		0.94	0.81 – 1.08	0.391	-----	----- -----	-------

Appendix 3.15 Continued

				Unadjusted			Adjusted		
	Number (%) Utilising Chemotherapy n = 1,912 (35.22)	Number (%) Not Utilising Chemotherapy n = 3,517 (64.78)	P Value[1]	OR	95% CI	P Value[2]	OR	95% CI	P Value[2]
ER Status			<0.001			**<0.001**			-------
Positive	890 (26.00)	2,533 (74.00)		1.00	----- -----	-------	-----	----- -----	-------
Negative	643 (64.43)	355 (35.57)		5.15	4.44 – 5.99	<0.001	-----	----- -----	-------
Unknown	379 (37.60)	629 (62.40)		1.71	1.48 -1.99	<0.001	-----	----- -----	-------
CCM (Between 78 to 6 Months Prior to Diagnosis)			<0.001			**<0.001**			**<0.001**
0	1,531 (42.98)	2,031 (57.02)		1.00	----- -----	-------	1.00	----- -----	-------
1-2	317 (24.61)	971 (75.39)		0.43	0.38 – 0.50	<0.001	0.44	0.38 – 0.51	<0.001
3+	64 (11.05)	515 (88.95)		0.16	0.13 – 0.22	<0.001	0.18	0.13 – 0.23	<0.001
Discussed at MDT			0.027			**0.0269**			**<0.001**
Yes	1,153 (36.63)	1,995 (63.37)		1.00	----- -----	-------	1.00	----- -----	-------
No	343 (32.39)	716 (67.61)		0.83	0.72 – 0.96	0.013	0.73	0.62 – 0.86	<0.001
Missing	416 (34.04)	806 (65.96)		0.89	0.78 – 1.03	0.110	0.84	0.72 – 0.97	0.020

[1]Chi-square P value
[2]P Values in bold are from likelihood ratio tests of the variable's contribution to the model. Unbolded P values are from a test of whether the OR is different from 1.
[3]For diagnosis year 2010, IMD_2010 was used and for diagnosis years 2013 - 2017, IMD_2015 was used.
[4]Other refers to Asian/British Asian, Black/African/Caribbean/Black British, mixed/multiple ethnic groups and other ethnic groups.
[5]Missing/unknown refers to missing and unknown ethnicity classifications.
[6]Unknown staging refers to missing and unstageable tumours.
[7]Other refers to undifferentiated or anaplastic, undetermined, and missing tumour grades.
CCM: Charlson Comorbidity Index; ER: Oestrogen receptor status; IMD: Index of multiple deprivation; MDT: Multidisciplinary team; OR: Odds ratio; 95% CI: 95% Confidence interval.

Appendix 3.16 Chemotherapy models (younger women, no surgery) for Chapter 3

Likelihood (OR with 95% CI and p values from logistic regression) of utilising chemotherapy in women by deprivation and adjusted for: diagnosis year, stage, ethnicity, comorbidities, and whether discussed at MDT if not receiving surgery and aged <60 with HER2+ breast cancer diagnosed between 01/01/2012 – 31/12/2017 (n = 4,626)

				Unadjusted			**Adjusted**		
	Number (%) Utilising Chemotherapy n = 3,778 (81.67)	Number (%) Not Utilising Chemotherapy n = 848 (18.33)	P Value[1]	OR	95% CI	P Value[2]	OR	95% CI	P Value[2]
IMD[3]			<0.001			**<0.001**			**<0.001**
1 (Least Deprived)	769 (75.61)	248 (24.39)		1.00	----- -----	-------	1.00	----- -----	-------
2	688 (77.92)	195 (22.08)		1.14	0.92 – 1.41	0.237	0.97	0.75 – 1.24	0.788
3	779 (86.08)	126 (13.92)		1.99	1.57 – 2.53	<0.001	1.71	1.31 – 2.24	<0.001
4	786 (83.26)	158 (16.74)		1.60	1.28 – 2.00	<0.001	1.25	0.97 – 1.62	0.089
5 (Most Deprived)	756 (86.20)	121 (13.80)		2.01	1.59 – 2.56	<0.001	1.60	1.21 – 2.11	0.001
Diagnosis Year			0.000			**0.0001**			**0.1065**
2012	327 (79.37)	85 (20.63)		0.68	0.51 – 0.90	0.008	0.73	0.53 – 1.01	0.061
2013	490 (80.59)	118 (19.41)		0.73	0.57 – 0.94	0.016	0.90	0.67 – 1.20	0.466
2014	490 (78.65)	133 (21.35)		0.65	0.51 – 0.83	0.001	0.73	0.54 – 0.97	0.028
2015	601 (77.85)	171 (22.15)		0.62	0.49 – 0.78	<0.001	0.74	0.57 – 0.97	0.031
2016	778 (84.02)	148 (15.98)		0.93	0.74 – 1.17	0.536	0.94	0.72 – 1.23	0.659
2017	1,092 (84.98)	193 (15.02)		1.00	----- -----	-------	1.00	----- -----	-------
Ethnicity			<0.001			**<0.001**			**<0.001**
White	3,064 (84.71)	553 (15.29)		1.00	----- -----	-------	1.00	----- -----	-------
Other Ethnic Group[4]	587 (81.87)	130 (18.13)		0.81	0.66 – 1.01	0.057	0.72	0.57 – 0.92	0.008
Missing/unknown[5]	127 (43.49)	165 (56.51)		0.14	0.11 – 0.18	<0.001	0.25	0.19 – 0.33	<0.001
Rural/Urban Indicator			<0.001			**0.001**			-------
Rural Village, Hamlet & Isolated Dwellings	291 (74.42)	100 (25.58)		0.65	0.51 – 0.84	0.001	-----	----- -----	-------
Rural Town & Fringe	337 (85.32)	58 (14.68)		1.30	0.96 – 1.76	0.085	-----	----- -----	-------
Urban City & Town	1,585 (82.38)	339 (17.62)		1.05	0.89 – 1.24	0.572	-----	----- -----	-------
Urban Conurbation	1,565 (81.68)	351 (18.32)		1.00	----- -----	-------	-----	----- -----	-------

Appendix 3.16 Continued

				Unadjusted			Adjusted		
	Number (%) Utilising Chemotherapy n = 3,778 (81.67)	Number (%) Not Utilising Chemotherapy n = 848 (18.33)	P Value[1]	OR	95% CI	P Value[2]	OR	95% CI	P Value[2]
Government Region			<0.001			**<0.001**			-------
North West	177 (85.51)	30 (14.49)		1.00	----- -----	-------	-----	----- -----	-------
North East	448 (80.58)	108 (19.42)		1.42	0.92 – 2.21	0.117	-----	----- -----	-------
West Midlands	419 (87.47)	60 (12.53)		1.68	1.20 – 2.37	0.003	-----	----- -----	-------
Yorkshire & the Humber	370 (89.16)	45 (10.84)		1.98	1.36 – 2.88	<0.001	-----	----- -----	-------
East Midlands	362 (85.99)	59 (14.01)		1.48	1.05 – 2.09	0.027	-----	----- -----	-------
East of England	473 (84.31)	88 (15.69)		1.30	0.95 – 1.77	0.101	-----	----- -----	-------
South East	571 (73.11)	210 (26.89)		0.66	0.50 – 0.85	0.002	-----	----- -----	-------
South West	294 (84.97)	52 (15.03)		1.36	0.95 – 1.96	0.094	-----	----- -----	-------
London	664 (77.21)	196 (22.79)		0.82	0.63 – 1.06	0.132	-----	----- -----	-------
Stage			<0.001			**<0.001**			**<0.001**
I	293 (67.05)	144 (32.95)		1.00	----- -----	-------	1.00	----- -----	-------
II	1,803 (88.99)	223 (11.01)		3.97	3.12 – 5.07	<0.001	3.70	2.81 – 4.87	<0.001
III	615 (90.98)	61 (9.02)		4.95	3.56 – 6.89	<0.001	4.36	3.03 – 6.27	<0.001
IV	640 (81.63)	144 (18.37)		2.18	1.67 – 2.86	<0.001	1.97	1.45 – 2.67	<0.001
Unknown[6]	427 (60.74)	276 (39.26)		0.76	0.59 – 0.98	0.032	0.98	0.73 – 1.32	0.897
Grade			<0.001			**<0.001**			-------
Well Differentiated (Low Grade)	42 (39.25)	65 (60.75)		0.11	0.08 – 0.17	<0.001	-----	----- -----	-------
Moderately Differentiated	1,420 (79.55)	365 (20.45)		0.68	0.58 – 0.80	<0.001	-----	----- -----	-------
Poorly Differentiated	2,170 (85.10)	380 (14.90)		1.00	----- -----	-------	-----	----- -----	-------
Other[7]	146 (79.35)	38 (20.65)		0.67	0.46 – 0.98	0.037	-----	----- -----	-------
Big Tumour Count			0.624			**0.626**			-------
1	3,511 (81.75)	784 (18.25)		1.00	----- -----	-------	-----	----- -----	-------
>1	267 (80.66)	64 (19.34)		0.93	0.70 – 1.24	0.624	-----	----- -----	-------

Appendix 3.16 Continued

	Number (%) Utilising Chemotherapy n = 3,778 (81.67)	Number (%) Not Utilising Chemotherapy n = 848 (18.33)	P Value[1]	Unadjusted OR	95% CI	P Value[2]	Adjusted OR	95% CI	P Value[2]
ER Status			<0.001			**<0.001**			-------
Positive	1,989 (79.62)	509 (20.38)		1.00	----- -----	-------	-----	----- -----	-------
Negative	1,136 (87.38)	164 (12.62)		1.77	1.47 – 2.14	<0.001	-----	----- -----	-------
Unknown	653 (78.86)	175 (21.14)		0.95	0.79 – 1.16	0.640	-----	----- -----	-------
CCM (Between 78 to 6 Months Prior to Diagnosis)			0.011			**0.0122**			**0.0035**
0	3,442 (81.43)	785 (18.57)		1.00	----- -----	-------	1.00	----- -----	-------
1-2	307 (85.99)	50 (14.01)		1.40	1.03 – 1.91	0.033	1.24	0.88 – 1.74	0.225
3+	29 (69.05)	13 (30.95)		0.51	0.26 – 0.98	0.044	0.29	0.14 – 0.61	0.001
Discussed at MDT			<0.001			**<0.001**			**<0.001**
Yes	2,278 (92.19)	193 (7.81)		1.00	----- -----	-------	1.00	----- -----	-------
No	635 (88.07)	86 (11.93)		0.63	0.48 – 0.82	0.001	0.60	0.45 – 0.80	<0.001
Missing	865 (60.32)	569 (39.68)		0.13	0.11 – 0.15	<0.001	0.16	0.13 – 0.19	<0.001

[1]Chi-square P value
[2]P Values in bold are from likelihood ratio tests of the variable's contribution to the model. Unbolded P values are from a test of whether the OR is different from 1.
[3]For diagnosis year 2012, IMD_2010 was used and for diagnosis years 2013 - 2017, IMD_2015 was used.
[4]Other refers to Asian/British Asian, Black/African/Caribbean/Black British, mixed/multiple ethnic groups and other ethnic groups.
[5]Missing/unknown refers to missing and unknown ethnicity classifications.
[6]Unknown staging refers to missing and unstageable tumours.
[7]Other refers to undifferentiated or anaplastic, undetermined, and missing tumour grades.
CCM: Charlson Comorbidity Index; ER: Oestrogen receptor status; IMD: Index of multiple deprivation; MDT: Multidisciplinary team; OR: Odds ratio; 95% CI: 95% Confidence interval.

Appendix 3.17 Chemotherapy models sensitivity analyses (older women, had surgery) for Chapter 3

				Unadjusted			Adjusted		
	Number (%) Utilising Chemotherapy	Number (%) Not Utilising Chemotherapy	P Value[1]	OR	95% CI	P Value[2]	OR	95% CI	P Value[2]
Full Analytical Cohort Original Model for Comparison (n = 15,202)									
IMD[3]	n = 7,236 (47.60)	7,966 (56.40)	0.715			**0.715**			**0.718**
1 (Least Deprived)	1,777 (47.96)	1,928 (52.04)		1.00	----- -----	-------	1.00	----- -----	-------
2	1,732 (48.12)	1,867 (51.88)		1.00	0.92 – 1.10	0.890	1.02	0.93 – 1.13	0.627
3	1,500 (47.86)	1,634 (52.14)		1.00	0.91 – 1.10	0.934	1.02	0.92 – 1.12	0.729
4	1,233 (46.88)	1,397 (53.12)		0.96	0.87 – 1.06	0.396	0.98	0.88 – 1.09	0.675
5 (Most Deprived)	994 (46.58)	1,140 (53.42)		0.95	0.85 – 1.05	0.308	0.95	0.85 – 1.07	0.396
Diagnosis Date Post Mandated SACT Submission 01/04/2014 – 31/12/2017 (n = 10,591) Sensitivity Analysis 1									
IMD[3]	n = 4,873 (46.01)	n = 5,718 (53.99)	0.743			**0.743**			**0.844**
1 (Least Deprived)	1,194 (46.40)	1,379 (53.60)		1.00	----- -----	-------	1.00	----- -----	-------
2	1,173 (46.71)	1,338 (53.29)		1.01	0.91 – 1.13	0.825	1.03	0.92 – 1.15	0.642
3	1,013 (46.13)	1,183 (53.87)		0.99	0.88 – 1.11	0.849	1.02	0.90 – 1.14	0.791
4	835 (45.41)	1,004 (54.59)		0.96	0.85 – 1.08	0.511	1.00	0.88 – 1.13	0.996
5 (Most Deprived)	658 (44.70)	814 (55.30)		0.93	0.82 – 1.06	0.295	0.95	0.83 – 1.09	0.457
Positive HER2+ Status Definition (n = 9,827) Sensitivity Analysis 2									
IMD[3]	n = 5,828 (59.31)	n = 3,999 (40.69)	0.625			**0.625**			**0.773**
1 (Least Deprived)	1,443 (59.88)	967 (40.12)		1.00	----- -----	-------	1.00	----- -----	-------
2	1,411 (60.35)	927 (39.65)		1.02	0.91 – 1.15	0.738	1.02	0.91 – 1.15	0.693
3	1,182 (58.66)	833 (41.34)		0.95	0.84 – 1.07	0.412	0.95	0.84 – 1.08	0.458
4	994 (58.61)	702 (41.39)		0.95	0.84 – 1.08	0.416	0.96	0.84 – 1.09	0.516
5 (Most Deprived)	798 (58.33)	570 (41.67)		0.94	0.82 – 1.07	0.354	0.97	0.84 – 1.11	0.639

[1]Chi-square P value; [2]Bolded P Values in are from likelihood ratio tests of the variable contribution to the model. Unbolded P values are from a test of whether the OR is different from 1; [3]For diagnosis year 2012, IMD_2010 was used and for diagnosis years 2013 - 2017, IMD_2015 was used.
All models adjusted for diagnosis year, ethnicity, stage, grade, comorbidities and whether discussed at MDT.

Appendix 3.18 Chemotherapy models sensitivity analyses (younger women, had surgery) for Chapter 3

				Unadjusted			Adjusted		
	Number (%) Utilising Chemotherapy	Number (%) Not Utilising Chemotherapy	P Value[1]	OR	95% CI	P Value[2]	OR	95% CI	P Value[2]
Full Analytical Cohort Original Model for Comparison (n = 14,922)									
IMD[3]	n = 11,518 (77.19)	n = 3,404 (22.81)	<0.001			**<0.001**			**0.001**
1 (Least Deprived)	2,488 (73.83)	882 (26.17)		1.00	----- -----		1.00	----- -----	-------
2	2,564 (76.56)	785 (23.44)		1.16	1.04 – 1.29	0.010	1.13	1.00 – 1.27	0.054
3	2,299 (77.28)	676 (22.72)		1.21	1.07 – 1.35	0.001	1.12	0.99 – 1.27	0.071
4	2,163 (78.17)	604 (21.83)		1.27	1.13 – 1.43	<0.001	1.14	1.00 – 1.29	0.054
5 (Most Deprived)	2,004 (81.43)	457 (18.57)		1.55	1.37 – 1.77	<0.001	1.36	1.19 – 1.57	<0.001
Diagnosis Date Post Mandated SACT Submission 01/04/2014 – 31/12/2017 (n = 10,003) Sensitivity Analysis 1									
IMD[3]	n = 7,441 (74.39)	n = 2,562 (25.61)	<0.001			**<0.001**			**0.001**
1 (Least Deprived)	1,603 (71.34)	644 (28.66)		1.00	----- -----	-------	1.00	----- -----	-------
2	1,680 (74.17)	585 (25.83)		1.15	1.01 – 1.32	0.033	1.09	0.94 – 1.26	0.238
3	1,440 (73.02)	532 (26.98)		1.09	0.95 – 1.24	0.224	0.98	0.84 – 1.14	0.780
4	1,392 (74.80)	469 (25.20)		1.19	1.04 – 1.37	0.013	1.02	0.88 – 1.19	0.791
5 (Most Deprived)	1,326 (79.98)	332 (20.02)		1.60	1.38 – 1.87	<0.001	1.35	1.14 – 1.59	<0.001
Positive HER2+ Status Definition (n = 10,839) Sensitivity Analysis 2									
IMD[3]	n = 9,397 (86.70)	n = 1,442 (13.30)	<0.001			**<0.001**			**<0.001**
1 (Least Deprived)	2,061 (83.58)	405 (16.42)		1.00	----- -----	-------	1.00	----- -----	-------
2	2,102 (85.45)	358 (14.55)		1.15	0.99 – 1.35	0.070	1.10	0.93 – 1.30	0.269
3	1,885 (88.00)	257 (12.00)		1.44	1.22 – 1.71	<0.001	1.32	1.11 – 1.59	0.002
4	1,770 (87.32)	257 (12.68)		1.35	1.14 – 1.60	<0.001	1.22	1.02 – 1.46	0.030
5 (Most Deprived)	1,579 (90.54)	165 (9.46)		1.88	1.55 – 2.28	<0.001	1.62	1.32 – 1.99	<0.001

[1]Chi-square P value; [2]Bolded P Values in are from likelihood ratio tests of the variable contribution to the model. Unbolded P values are from a test of whether the OR is different from 1; [3]For diagnosis year 2012, IMD_2010 was used and for diagnosis years 2013 - 2017, IMD_2015 was used.
All models adjusted for: diagnosis year, ethnicity, stage, grade, and whether discussed at MDT. Abbreviations: HER2: Human epidermal growth factor: IMD: Index of multiple deprivation; OR: Odds ratio; SACT: Systemic Anti-Cancer Therapy: 95% CI: 95% Confidence interval.

Appendix 3.19 Chemotherapy models sensitivity analyses (older women, no surgery) for Chapter 3

				Unadjusted			Adjusted		
	Number (%) Utilising Chemotherapy	Number (%) Not Utilising Chemotherapy	P Value[1]	OR	95% CI	P Value[2]	OR	95% CI	P Value[2]
Full Analytical Cohort Original Model for Comparison (n = 5,429)									
IMD[3]	n = 1,912 (35.22)	n = 3,517 (64.78)	0.013			**0.013**			**0.006**
1 (Least Deprived)	460 (38.98)	720 (61.02)		1.00	----- -----	-------	1.00	----- -----	-------
2	403 (35.48)	733 (64.52)		0.86	0.73 – 1.02	0.081	0.81	0.68 – 0.97	0.025
3	371 (33.33)	742 (66.67)		0.78	0.66 – 0.93	0.005	0.77	0.64 – 0.92	0.005
4	335 (32.40)	699 (67.60)		0.75	0.63 – 0.89	0.001	0.71	0.59 – 0.86	<0.001
5 (Most Deprived)	343 (35.51)	623 (64.49)		0.86	0.72 – 1.03	0.098	0.85	0.70 – 1.02	0.087
Diagnosis Date Post Mandated SACT Submission 01/04/2014 – 31/12/2017 (n = 4,085) Sensitivity Analysis 1									
IMD[3]	n = 1,448 (35.45)	n = 2,637 (64.55)	0.007			**0.007**			**0.014**
1 (Least Deprived)	354 (40.18)	527 (59.82)		1.00	----- -----		1.00	----- -----	
2	316 (36.03)	561 (63.97)		0.84	0.69 – 1.02	0.073	0.83	0.67 – 1.02	0.072
3	278 (33.21)	559 (66.79)		0.74	0.61 – 0.90	0.003	0.75	0.61 – 0.93	0.009
4	255 (32.28)	535 (67.72)		0.71	0.58 – 0.87	0.001	0.69	0.56 – 0.86	0.001
5 (Most Deprived)	245 (35.00)	455 (65.00)		0.80	0.65 – 0.98	0.035	0.85	0.68 – 1.06	0.146
Positive HER2+ Status Definition (n = 3,335) Sensitivity Analysis 2									
IMD[3]	n = 1,544 (46.30)	n = 1,791 (53.70)	0.204			**0.205**			**0.069**
1 (Least Deprived)	372 (49.87)	374 (50.13)		1.00	----- -----	-------	1.00	----- -----	-------
2	324 (46.09)	379 (53.91)		0.86	0.70 – 1.06	0.150	0.78	0.62 – 0.97	0.025
3	304 (44.44)	380 (55.56)		0.80	0.65 – 0.99	0.040	0.77	0.62 – 0.97	0.023
4	271 (44.14)	343 (55.86)		0.79	0.64 – 0.98	0.035	0.74	0.59 – 0.94	0.013
5 (Most Deprived)	273 (46.43)	315 (53.57)		0.87	0.70 – 1.08	0.212	0.86	0.68 – 1.09	0.200

[1]Chi-square P value; [2]Bolded P Values in are from likelihood ratio tests of the variable contribution to the model. Unbolded P values are from a test of whether the OR is different from 1; [3]For diagnosis year 2012, IMD_2010 was used and for diagnosis years 2013 - 2017, IMD_2015 was used.
All models adjusted for: diagnosis year, ethnicity, stage, grade, comorbidities, and whether discussed at MDT.

Appendix 3.20 Chemotherapy models sensitivity analyses (younger women, no surgery) for Chapter 3

				Unadjusted			Adjusted		
	Number (%) Utilising Chemotherapy	Number (%) Not Utilising Chemotherapy	P Value[1]	OR	95% CI	P Value[2]	OR	95% CI	P Value[2]
Full Analytical Cohort Original Model for Comparison (n = 4,626)									
IMD[3]	n = 3,778 (81.67)	n = 848 (18.33)	<0.001			**<0.001**			**<0.001**
1 (Least Deprived)	769 (75.61)	248 (24.39)		1.00	----- -----	-------	1.00	----- -----	-------
2	688 (77.92)	195 (22.08)		1.14	0.92 – 1.41	0.237	0.97	0.75 – 1.24	0.788
3	779 (86.08)	126 (13.92)		1.99	1.57 – 2.53	<0.001	1.71	1.31 – 2.24	<0.001
4	786 (83.26)	158 (16.74)		1.60	1.28 – 2.00	<0.001	1.25	0.97 – 1.62	0.089
5 (Most Deprived)	756 (86.20)	121 (13.80)		2.01	1.59 – 2.56	<0.001	1.60	1.21 – 2.11	0.001
Diagnosis Date Post Mandated SACT Submission 01/04/2014 – 31/12/2017 (n = 3,467) Sensitivity Analysis 1									
IMD[3]	n = 2,849 (82.17)	n = 618 (17.83)	<0.001			**<0.001**			**0.002**
1 (Least Deprived)	607 (77.32)	178 (22.68)		1.00	----- -----	-------	1.00	----- -----	-------
2	506 (77.85)	144 (22.15)		1.03	0.80 – 1.32	0.814	0.90	0.67 – 1.20	0.464
3	564 (85.45)	96 (14.55)		1.72	1.31 – 2.26	<0.001	1.45	1.06 – 2.00	0.021
4	603 (84.10)	114 (15.90)		1.55	1.20 – 2.01	0.001	1.20	0.88 – 1.63	0.240
5 (Most Deprived)	569 (86.87)	86 (13.13)		1.94	1.46 – 2.57	<0.001	1.62	1.16 – 2.26	0.004
			<0.001			**<0.001**			**0.002**
Positive HER2+ Status Definition (n = 3,711) Sensitivity Analysis 2									
IMD[3]	n = 3,142 (11.34)	n = 569 (2.05)	<0.001			**<0.001**			**<0.001**
1 (Least Deprived)	628 (77.63)	181 (22.37)		1.00	----- -----	-------	1.00	----- -----	-------
2	578 (80.95)	136 (19.05)		1.22	0.95 – 1.57	0.111	1.06	0.79 – 1.43	0.694
3	651 (89.67)	75 (10.33)		2.50	1.87 – 3.35	<0.001	2.27	1.62 – 3.17	<0.001
4	645 (86.23)	103 (13.77)		1.80	1.38 – 2.34	<0.001	1.49	1.09 – 2.03	0.013
5 (Most Deprived)	640 (89.64)	74 (10.36)		2.49	1.86 – 3.34	<0.001	2.08	1.48 – 2.92	<0.001

[1]Chi-square P value; [2]Bolded P Values in are from likelihood ratio tests of the variable contribution to the model. Unbolded P values are from a test of whether the OR is different from 1; [3]For diagnosis year 2012, IMD_2010 was used and for diagnosis years 2013 - 2017, IMD_2015 was used.
All models adjusted for: diagnosis year, ethnicity, stage, grade, comorbidities, and whether discussed at MDT.

Appendix 3.21 Route to diagnosis for the HER2+ breast cancer cohort presented by date of diagnosis for Chapter 3*

Date	Route to Diagnosis N (%)					
	Screening	Emergency[1]	2WW[2]	Other Route[3]	Unknown	Missing
2012	1,067 (22.63)	152 (3.22)	2,797 (59.31)	556 (11.79)	142 (3.01)	2 (0.04)
2013	1,364 (23.74)	162 (2.82)	3,404 (59.25)	629 (10.95)	182 (3.17)	4 (0.07)
2014	1,559 (25.07)	175 (2.81)	3,580 (57.57)	695 (11.18)	203 (3.26)	7 (0.11)
2015	1,699 (24.34)	198 (2.84)	4,107 (58.84)	733 (10.50)	241 (3.45)	2 (0.03)
2016	1,785 (22.59)	231 (2.92)	4,614 (58.39)	905 (11.45)	357 (4.52)	10 (0.13)

***Route to diagnosis data was not available for 2017.**

[1]Emergency refers to A&E, emergency GP referral, emergency transfer and emergency admission or attendance

[2]2WW refers to the urgent care pathway (target of 14 days from the point of referral for suspected cancer. symptoms to the point of first assessment with a specialist at a hospital).

[3]Other route refers to GP referral, other outpatient, and inpatient elective.

Abbreviations: N: Number; 2WW: Two week wait.

Appendix 3.22 HER2+ classification for the cohort of interest presented by date of diagnosis for Chapter 3

Date	HER2 Classification (Number)		
	B[1]	P[2]	Pm[3]
2012	469	4,247	0
2013	853	4,890	2
2014	1,290	4,891	38
2015	2,113	4,569	298
2016	3,010	4,622	270
2017	3,936	4,493	188

[1]B refers to borderline HER2 classification.
[2]P refers to positive HER2 classification.
[3]Pm refers to a classification whereby the patient has one positive test but may have had a different result in other tests (one record as P but at least one record with something else).
Abbreviations: HER2: Human epidermal growth factor receptor 2.

Appendix 4.1 Novel anti-cancer therapy utilisation by stage at initial diagnosis in the NSCLC population for Chapter 4

Stage at Diagnosis	Utilised a Novel-Anti Cancer Therapy Number (%)
I	384 (3.90)
II	437 (4.43)
III	2,307 (23.41)
IV	6,545 (66.42)
Unknown	181 (1.84)

Appendix 4.2 Any Novel anti-cancer therapy SACT references used for coding as any novel cancer therapy for Chapter 4

SACT Drug Classification	SACT Analysis Group Listings	SACT Benchmark Group Listings	SACT Drug Group Listings
Any Novel Anti-Cancer Therapy	**AFATINIB** CARBO + **GEFITINIB** + PEMETREXED **ERLOTINIB** **GEFITINIB** IDEAL TRIAL **OSIMERTINIB** TIMELY TRIAL **ALECTINIB** **BRIGATINIB** **CERITINIB** **CRIZOTINIB** **LORLATINIB** **DABRAFENIB** **DABRAFENIB** + **TRAMETINIB** **ATEZOLIZUMAB** CARBOPLATIN + **PEMBROLIZUMAB** + PEMETREXED CISPLATIN + **PEMBROLIZUMAB** + PEMETREXED **DURVALUMAB** **IPILIMUMAB** + **NIVOLUMAB** **NIVOLUMAB** **PEMBROLIZUMAB** **PEMBROLIZUMAB** + PEMETREXED **BEVACIZUMAB** **BEVACIZUMAB** + CARBOPLATIN + PACLITAXEL **BEVACIZUMAB** + CISPLATIN + PACLITAXEL **BEVACIZUMAB** + PACLITAXEL DOCETAXEL + **NINTEDANIB** **NINTEDANIB** **ATEZOLIZUMAB** + **BEVACIZUMAB** **ATEZOLIZUMAB** + **BEVACIZUMAB** + CARBOPLATIN + PACLITAXEL MATRIX TRIAL TAXTORC TRIAL	**AFATINIB** CARBO + **GEFITINIB** + PEMETREXED **ERLOTINIB** **GEFITINIB** IDEAL TRIAL **OSIMERTINIB** TIMELY TRIAL **ALECTINIB** **BRIGATINIB** **CERITINIB** **CRIZOTINIB** **LORLATINIB** **DABRAFENIB** **DABRAFENIB** + **TRAMETINIB** **ATEZOLIZUMAB** CARBOPLATIN + **PEMBROLIZUMAB** + PEMETREXED CISPLATIN + **PEMBROLIZUMAB** + PEMETREXED **DURVALUMAB** **IPILIMUMAB** + **NIVOLUMAB** **NIVOLUMAB** **PEMBROLIZUMAB** **PEMBROLIZUMAB** + PEMETREXED **BEVACIZUMAB** **BEVACIZUMAB** + CARBOPLATIN + PACLITAXEL **BEVACIZUMAB** + CISPLATIN + PACLITAXEL **BEVACIZUMAB** + PACLITAXEL DOCETAXEL + **NINTEDANIB** **NINTEDANIB** **ATEZOLIZUMAB** + **BEVACIZUMAB** **ATEZOLIZUMAB** + **BEVACIZUMAB** + CARBOPLATIN + PACLITAXEL MATRIX TRIAL TAXTORC TRIAL	**AFATINIB** **AZD9291** **ERLOTINIB** **GEFITINIB** **OSIMERTINIB** **ROCILETINIB** **ALECTINIB** **BRIGATINIB** **CERITINIB** **CRIZOTINIB** **LORLATINIB** **X-396** **CABOZANTINIB** **CAPMATINIB** **DABRAFENIB** **ENTRECTINIB** **TRAMETINIB** **ATEZOLIZUMAB** **BMS-986016** **DURVALUMAB** **IPILIMUMAB** **MED14736** **MEDI4736** **NIVOLUMAB** **NKTR-214** **PEMBROLIZUMAB** **PDR001** **TSR-042** **TREMELIMUMAB** **BEVACIZUMAB** **NINTEDANIB** **RAMUCIRUMAB** **AZD5363** **BGB324** **BKM120** **GANETESPIB** **LCL161** **MSB0011359C** **ONARTUZUMAB** **VELIPARIB** **VISTUSERTIB**

Appendix 4.2 Continued

*Trials listed in table only if it is certain that all patients received that drug.
Bold text indicates the reference to a novel anti-cancer therapy within the drug regimen.

Novel Anti-Cancer Therapy Trial Information

Ideal Trial: Randomised, double blind parallel group phase II trial of gefitinib for previously treated with one or two chemotherapy regimens (at least one containing platinum) with advanced NSCLC were assigned either 250mg or 500mg oral OD dose.
Matrix Trial: Parallel, multi-centre single arm phase II trial each testing an experimental targeted drug in a population stratified by multiple pre-specified actionable target putative biomarkers. Arm A: AZD4547 (FGFR inhibitor); Arm B: vistusertib; Arm C: palbociclib; Arm D: crizotinib; Arm E: selumetinib & docetaxel; Arm Fl AZD5363 (AKT inhibitor); Arm G: osimertinib; Arm H: sitravatinib; Arm J: AZD6738 (ART inhibitor) & durvalumab; Arm NA (if have no gene change): durvalumab; Arm NAJ (if have no gene change): AZD6738 (ART inhibitor) & durvalumab.
Taxtorc Trial: Phase I trial looking at use of vistusertib and paclitaxel for solid tumours including ovarian cancer and NSCLC.
Timely Trial: Phase II trial looking at afatinib for NSCLC in patients who cannot have chemotherapy.

Pre-clinical Drug Names

AZD5363: Capivasrtib; AZD929: Osimertinib; BGB324: Bemcentinib; BKM120: Buparlisib; BMS-986016: Relatilimab; LCL161: Ipasertib; MED14736: Durvalumab; MEDI4736: Durvalumab; MSB0011359C: Bifunctional fusion protein targeting PD-L1 and TGF-beta. No name; NKTR-214: Bempegaldesleukin; PDR001: Spartalizumab; TSR-042: Dostarlimab; X-396: Ensartinib.

Appendix 4.3 Exploratory analysis any novel anti-cancer therapy sub-group classifications SACT references used for coding for Chapter 4

SACT Drug Classification	SACT Analysis Group Listings	SACT Benchmark Group Listings	SACT Drug Group Listings
Any Targeted Therapy	AFATINIB CARBO + GEFITINIB + PEMETREXED ERLOTINIB GEFITINIB IDEAL TRIAL OSIMERTINIB TIMELY TRIAL ALECTINIB BRIGATINIB CERITINIB CRIZOTINIB LORLATINIB DABRAFENIB DABRAFENIB + TRAMETINIB	AFATINIB CARBO + GEFITINIB + PEMETREXED ERLOTINIB GEFITINIB IDEAL TRIAL OSIMERTINIB TIMELY TRIAL ALECTINIB BRIGATINIB CERITINIB CRIZOTINIB LORLATINIB DABRAFENIB DABRAFENIB + TRAMETINIB	AFATINIB AZD9291 ERLOTINIB GEFITINIB OSIMERTINIB ROCILETINIB ALECTINIB BRIGATINIB CERITINIB CRIZOTINIB LORLATINIB X-396 CABOZANTINIB CAPMATINIB DABRAFENIB ENTRECTINIB TRAMETINIB
Targeted Therapy EGFRi	AFATINIB CARBO + GEFITINIB + PEMETREXED ERLOTINIB GEFITINIB IDEAL TRIAL OSIMERTINIB TIMELY TRIAL	AFATINIB CARBO + GEFITINIB + PEMETREXED ERLOTINIB GEFITINIB IDEAL TRIAL OSIMERTINIB TIMELY TRIAL	AFATINIB AZD9291 ERLOTINIB GEFITINIB OSIMERTINIB ROCILETINIB
Targeted Therapy EGFRi (No Erlotinib)	AFATINIB CARBO + GEFITINIB + PEMETREXED GEFITINIB IDEAL TRIAL OSIMERTINIB TIMELY TRIAL	AFATINIB CARBO + GEFITINIB + PEMETREXED GEFITINIB IDEAL TRIAL OSIMERTINIB TIMELY TRIAL	AFATINIB AZD9291 GEFITINIB OSIMERTINIB ROCILETINIB
Targeted Therapy ALKi	ALECTINIB BRIGATINIB CERITINIB CRIZOTINIB LORLATINIB	ALECTINIB BRIGATINIB CERITINIB CRIZOTINIB LORLATINIB	ALECTINIB BRIGATINIB CERITINIB CRIZOTINIB LORLATINIB X-396
Biologicals (with Mixed Bevacizumab Classification Variables)	ATEZOLIZUMAB + BEVACIZUMAB ATEZOLIZUMAB + BEVACIZUMAB + CARBOPLATIN + PACLITAXEL BEVACIZUMAB BEVACIZUMAB + CARBOPLATIN + PACLITAXEL BEVACIZUMAB + CISPLATIN + PACLITAXEL BEVACIZUMAB + PACLITAXEL DOCETAXEL + NINTEDANIB NINTEDANIB	ATEZOLIZUMAB + BEVACIZUMAB ATEZOLIZUMAB + BEVACIZUMAB + CARBOPLATIN + PACLITAXEL BEVACIZUMAB BEVACIZUMAB + CARBOPLATIN + PACLITAXEL BEVACIZUMAB + CISPLATIN + PACLITAXEL BEVACIZUMAB + PACLITAXEL DOCETAXEL + NINTEDANIB NINTEDANIB	BEVACIZUMAB NINTEDANIB RAMUCIRUMAB

Appendix 4.3 Continued

SACT Drug Classification	SACT Analysis Group Listings	SACT Benchmark Group Listings	SACT Drug Group Listings
Immunotherapy (with Mixed Atezolizumab Classification Variables)	ATEZOLIZUMAB ATEZOLIZUMAB + BEVACIZUMAB ATEZOLIZUMAB + BEVACIZUMAB + CARBOPLATIN + PACLITAXEL CARBOPLATIN + PEMBROLIZUMAB + PEMETREXED CISPLATIN + PEMBROLIZUMAB + PEMETREXED DURVALUMAB IPILIMUMAB + NIVOLUMAB NIVOLUMAB PEMBROLIZUMAB PEMBROLIZUMAB + PEMETREXED	ATEZOLIZUMAB ATEZOLIZUMAB + BEVACIZUMAB ATEZOLIZUMAB + BEVACIZUMAB + CARBOPLATIN + PACLITAXEL CARBOPLATIN + PEMBROLIZUMAB + PEMETREXED CISPLATIN + PEMBROLIZUMAB + PEMETREXED DURVALUMAB IPILIMUMAB + NIVOLUMAB NIVOLUMAB PEMBROLIZUMAB PEMBROLIZUMAB + PEMETREXED	ATEZOLIZUMAB BMS-986016 DURVALUMAB IPILIMUMAB MED14736 MEDI4736 NIVOLUMAB NKTR-214 PEMBROLIZUMAB PDR001 TSR-042 TREMELIMUMAB

*Trials listed in table only if it is certain that all patients received that drug.
Bold text indicates the reference to a novel anti-cancer therapy within the drug regimen.

Novel Anti-Cancer Therapy Trial Information

Ideal Trial: Randomised, double blind parallel group phase II trial of gefitinib for previously treated with one or two chemotherapy regimens (at least one containing platinum) with advanced NSCLC were assigned either 250mg or 500mg oral OD dose.
Matrix Trial: Parallel, multi-centre single arm phase II trial each testing an experimental targeted drug in a population stratified by multiple pre-specified actionable target putative biomarkers. Arm A: AZD4547 (FGFR inhibitor); Arm B: vistusertib; Arm C: palbociclib; Arm D: crizotinib; Arm E: selumetinib & docetaxel; Arm Fl AZD5363 (AKT inhibitor); Arm G: osimertinib; Arm H: sitravatinib; Arm J: AZD6738 (ART inhibitor) & durvalumab; Arm NA (if have no gene change): durvalumab; Arm NAJ (if have no gene change): AZD6738 (ART inhibitor) & durvalumab.
Taxtorc Trial: Phase I trial looking at use of vistusertib and paclitaxel for solid tumours including ovarian cancer and NSCLC.
Timely Trial: Phase II trial looking at afatinib for NSCLC in patients who cannot have chemotherapy.

Pre-clinical Drug Names

AZD5363: Capivasrtib; AZD929: Osimertinib; BGB324: Bemcentinib; BKM120: Buparlisib; BMS-986016: Relatilimab; LCL161: Ipasertib; MED14736: Durvalumab; MEDI4736: Durvalumab; MSB0011359C: Bifunctional fusion protein targeting PD-L1 and TGF-beta. No name; NKTR-214: Bempegaldesleukin; PDR001: Spartalizumab; TSR-042: Dostarlimab; X-396: Ensartinib.

Appendix 4.4 Demographic and clinical characteristics of a cohort with a first invasive primary Stage IV NSCLC diagnosed between 01/01/2012 – 31/12/2017 who had a SACT record (n = 90,785) for Chapter 4

	N (%)
IMD[1]	
1 (Least Deprived)	13,193 (14.53)
2	16,579 (18.26)
3	18,451 (20.32)
4	20,105 (22.15)
5 (Most Deprived)	22,457 (24.74)
Sex	
Male	50,011 (55.09)
Female	40,774 (44.91)
Age (Years)	
<50	2,466 (2.72)
50 – 59	8,802 (9.70)
60 – 69	23,502 (25.89)
70 – 79	30,762 (33.88)
80 – 85	21,337 (23.50)
90+	3,916 (4.31)
Diagnosis Year	
2012	14,952 (16.47)
2013	14,731 (16.23)
2014	15,102 (16.63)
2015	15,522 (17.10)
2016	15,506 (17.08)
2017	14,972 (16.49)
Ethnicity	
White	83,469 (91.94)
Other Ethnic Group[2]	3,558 (3.92)
Missing/unknown[3]	3,758 (4.14)
Rural/Urban Indicator	
Rural Village, Hamlet & Isolated Dwellings	7,362 (8.11)
Rural Town & Fringe	8,898 (9.80)
Urban City & Town	40,516 (44.63)
Urban Conurbation	34,009 (37.46)
Histology	
NSCLC NOS	39,271 (43.26)
NSCLC Adenocarcinoma	36,541 (40.25)
NSCLC Squamous	14,120 (15.55)
NSCLC Large Cell	853 (0.94)
Big Tumour Count	
1	77,683 (85.57)
>1	13,102 (14.43)

Appendix 4.4 Continued

	N (%)
IMD[1]	
CCM (Between 78 to 6 Months Prior to Diagnosis)	
0	54,963 (60.54)
1 – 2	25,540 (28.13)
3+	10,282 (11.33)
SACT Record	
Yes (In Time Range)[4]	21,441 (23.62)
Yes (Not in Time Range)[4]	3,186 (3.51)
No	66,158 (72.87)
Treatment	
Utilised Chemotherapy	26,704 (29.41)
Utilised Surgery	1,261 (1.39)
Utilised Radiotherapy	25,414 (27.99)
Utilised a Novel Anti-Cancer Therapy	6,545 (7.21)
Utilised a Targeted Therapy	3,484 (3.84)
Utilised a Biological	746 (0.82)
Utilised an Immunotherapy	2,604 (2.87)

[1]For diagnosis year 2012, IMD_2010 was used and for diagnosis years 2013 - 2017, IMD_2015 was used.
[2]Other ethnic group refers to Asian/British Asian, Black/African/Caribbean/Black British, mixed/multiple ethnic groups and other ethnic groups.
[3]Missing/unknown ethnicity refers to unknown and missing ethnicity classifications.
[4]In time range refers to an SACT record up to 56 days before and 2 years post diagnosis.
CCM, Charlson comorbidity index; IMD: Index of Multiple Deprivation; NOS, Not Otherwise Specified; NSCLC, Non-Small Cell Lung Cancer; SACT; Systemic Anti-Cancer Therapy Dataset.

Appendix 4.5 Demographic and clinical characteristics of a cohort with a first invasive primary NSCLC diagnosed between 01/01/2012 - 31/12/2017 who had a SACT record (n = 50,044), had a SACT record in the time range (n = 41,998) and who did not have a SACT record (n = 145,345) for Chapter 4

Characteristic	Had a SACT Record n = 50,044	Had a SACT Record in the Time Range[1] n = 41,998	Did not Have a SACT Record n = 145,343
		Number (%)	
IMD[2]			
1 (Least Deprived)	7,881 (15.75)	6,719 (16.00)	19,753 (13.59)
2	9,475 (18.93)	7,972 (18.98)	25,723 (17.70)
3	10.093 (20.17)	8,457 (20.14)	28,914 (19.89)
4	10.704 (21.39)	8,967 (21.35)	32,956 (22.67)
5 (Most Deprived)	11,891 (23.76)	9,883 (23.53)	37,997 (26.14)
Sex			
Male	27,532 (55.02)	23,064 (54.92)	78,185 (53.79)
Female	22,512 (44.98)	18,934 (45.08)	67,158 (46.21)
Age (Years)			
<50	2,347 (4.69)	2,053 (4.89)	1,859 (1.28)
50 – 59	7,966 (15.92)	6,811 (16.22)	8,643 (7.23)
60 – 69	19,145 (38.26)	16,138 (38.43)	29,416 (20.24)
70 – 79	17,275 (34.52)	14,372 (34.22)	51,019 (35.10)
80 – 85	3,224 (6.44)	2,572 (6.12)	45,333 (31.19)
90+	87 (0.17)	52 (0.12)	9,073 (6.24)
Diagnosis Year			
2012	6,792 (13.57)	4,263 (10.15)	25,338 (17.43)
2013	8,137 (16.26)	5,978 (14.23)	24,291 (16.71)
2014	8,586 (17.16)	7,314 (17.42)	23,975 (16.50)
2015	8,796 (17.58)	7,919 (18.86)	23,873 (16.43)
2016	8,780 (17.54)	8,094 (19.27)	24,206 (16.65)
2017	8,953 (17.89)	8,430 (20.07)	23,660 (16.28)
Ethnicity			
White	46,621 (93.16)	38,979 (92.81)	133,069 (91.56)
Other Ethnic Group[3]	2,413 (4.82)	2,109 (5.02)	4,594 (3.16)
Missing/unknown[4]	1,010 (2.02)	910 (2.17)	7,680 (5.28)
Rural/Urban Indicator			
Rural Village, Hamlet & Isolated Dwellings	4,377 (8.75)	3,662 (8.72)	11,007 (7.57)
Rural Town & Fringe	4,828 (9.65)	4,087 (9.73)	13,970 (9.61)
Urban City & Town	21,362 (42.69)	17,768 (42.31)	65,143 (44.82)
Urban Conurbation	19,477 (38.92)	16,481 (39.24)	55,223 (37.99)
Stage			
I	4,308 (8.61)	2,662 (6.34)	28,100 (19.33)
II	4,892 (9.78)	3,985 (9.49)	10,429 (7.18)
III	14,847 (29.67)	12,914 (30.75)	21,995 (15.13)
IV	24,627 (49.21)	21,441 (51.05)	66,158 (45.52)
Unknown[5]	1,370 (2.74)	996 (2.37)	18,661 (12.84)

Appendix 4.5 Continued

Characteristic	Had a SACT Record	Had a SACT Record in the Time Range	Did not Have a SACT Record
	n = 50,044	n = 41,998	n = 145,343
	Number (%)		
Histology			
NSCLC NOS	7,362 (14.71)	5,866 (13.97)	75,591 (52.01)
NSCLC Adenocarcinoma	27,166 (54.28)	23,039 (54.86)	41,943 (28.86)
NSCLC Squamous	14,731 (29.44)	12,427 (29.59)	26,803 (18.44)
NSCLC Large Cell	785 (1.57)	666 (1.59)	1,006 (0.69)
Big Tumour Count			
1	39,595 (79.12)	34,201 (81.43)	119,535 (82.24)
>1	10,449 (20.88)	7,797 (18.57)	25,808 (17.76)
CCM (Between 78 to 6 Months Prior to Diagnosis)			
0	33,725 (67.39)	29,008 (69.07)	73,145 (50.33)
1 – 2	12,656 (25.29)	10,369 (24.69)	48,368 (33.28)
3+	3,663 (7.32)	2,621 (6.24)	23,830 (16.40)
SACT Record			
Yes (In Time Range)[1]	41,998 (83.92)	41,998 (100.00)	----------
Yes (Not in Time Range)[1]	8,046 (16.08)	----------	----------
Treatment			
Utilised Chemotherapy	43,878 (87.68)	38,686 (92.11)	5,395 (3.71)
Utilised Surgery	9,903 (19.79)	7,418 (17.66)	19,936 (13.72)
Utilised Radiotherapy	19,437 (38.84)	16,668 (39.69)	34,398 (23.67)
Utilised Any Novel Anti-Cancer Therapy	9,854 (19.69)	9,854 (23.46)	----------
Utilised a Targeted Therapy	4,783 (9.56)	4,783 (11.39)	----------
Utilised a Biological	1,039 (2.08)	1,039 (2.47)	----------
Utilised an Immunotherapy	4,398 (8.79)	4,398 (10.47)	----------

[1]In time range refers to an SACT record up to 56 days before and 2 years post diagnosis.
[2]For diagnosis year 2012, IMD_2010 was used and for diagnosis years 2013 - 2017, IMD_2015 was used.
[3]Other ethnic group refers to Asian/British Asian, Black/African/Caribbean/Black British, mixed/multiple ethnic groups and other ethnic groups.
[4]Missing/unknown refers to unknown and missing ethnicity classifications.
[5]Unknown staging refers to missing and unstageable tumours.
Abbreviations: CCM: Charlson Comorbidity Index; IMD: Index of multiple deprivation; NOS: Not otherwise specified; NSCLC: Non-small cell lung cancer; SACT: Systemic Anti-Cancer Therapy Dataset.

Appendix 4.6 Demographic and clinical characteristics by deprivation category for a cohort with a first primary invasive NSCLC cancer diagnosed between 01/01/2012 – 31/12/2017 (n = 195,387) for Chapter 4

	IMD 1[1] (Least Deprived) N (%)	IMD 2 N (%)	IMD 3 N (%)	IMD 4 N (%)	IMD 5 (Most Deprived) N (%)	P Value[2]
Sex						<0.001
Male	15,120 (54.72)	19,199 (54.55)	21,153 (54.23)	23,502 (53.83)	26,743 (53.61)	
Female	12,514 (45.28)	15,999 (45.45)	17,854 (45.77)	20,158 (46.17)	23,145 (46.39)	
Age (Years)						<0.001
<40	68 (0.25)	86 (0.24)	89 (0.23)	145 (0.33)	164 (0.33)	
40 – 49	408 (1.48)	503 (1.43)	664 (1.70)	884 (2.02)	1,195 (2.40)	
50 – 59	1,804 (6.53)	2,451 (6.96)	3,063 (7.85)	3,899 (8.93)	5,392 (10.81)	
60 – 69	6,319 (22.87)	8,323 (23.65)	9,446 (24.22)	10,957 (25.10)	13,516 (27.09)	
70 – 79	9,852 (35.65)	12,429 (35.31)	13,721 (35.18)	15,234 (34.89)	17,058 (34.19)	
80 – 85	7,609 (27.53)	9,512 (27.02)	10,091 (25.87)	10,539 (24.14)	10,806 (21.66)	
90+	1,574 (5.70)	1,894 (5.38)	1,933 (4.96)	2,002 (4.59)	1,757 (3.52)	
Diagnosis Year						<0.001
2012	4,351 (15.75)	5,883 (16.71)	6,515 (16.70)	7,380 (16.90)	8,001 (16.04)	
2013	4,451 (16.11)	5,814 (16.52)	6,570 (16.84)	7,194 (16.48)	8,399 (16.84)	
2014	4,736 (17.14)	5,783 (16.43)	6,388 (16.38)	7,274 (16.66)	8,380 (16.80)	
2015	4,654 (16.84)	5,878 (16.70)	6,478 (16.61)	7,207 (16.51)	8,452 (16.94)	
2016	4,740 (17.15)	5,933 (16.86)	6,484 (16.62)	7,403 (16.96)	8,426 (16.89)	
2017	4,702 (17.02)	5,907 (16.78)	6,572 (16.85)	7,202 (16.50)	8,230 (16.50)	
Ethnicity						<0.001
White	25,629 (92.74)	32,814 (93.23)	35,977 (92.23)	39,988 (91.59)	45,282 (90.77)	
Other Ethnic Group[3]	609 (2.20)	747 (2.12)	1,228 (3.15)	1,776 (4.07)	2,647 (5.31)	
Missing/unknown[4]	1,396 (5.05)	1,637 (4.65)	1,802 (4.62)	1,896 (4.34)	1,959 (3.93)	

Appendix 4.6 Continued

	IMD 1[1] (Least Deprived) N (%)	IMD 2 N (%)	IMD 3 N (%)	IMD 4 N (%)	IMD 5 (Most Deprived) N (%)	P Value[2]
Rural/Urban Indicator						<0.001
Rural Village, Hamlet & Isolated Dwellings	4,628 (16.75)	6,119 (17.38)	3,567 (9.14)	916 (2.10)	154 (0.31)	
Rural Town & Fringe	3,366 (12.18)	4,992 (14.18)	5,117 (13.12)	3,937 (9.02)	1,386 (2.78)	
Urban City & Town	12,778 (46.24)	15,098 (42.89)	18,318 (46.96)	20,930 (47.94)	19,381 (38.85)	
Urban Conurbation	6,862 (24.83)	8,989 (25.54)	12,005 (30.78)	17,877 (40.95)	28,967 (58.06)	
Stage						<0.001
I	4,601 (16.65)	5,661 (16.08)	6,229 (15.97)	7,176 (16.44)	8,741 (17.52)	
II	2,115 (7.65)	2,781 (7.90)	2,973 (7.62)	3,384 (7.75)	4,068 (8.15)	
III	4,832 (17.49)	6,476 (18.40)	7,223 (18.52)	8,445 (19.34)	9,866 (19.78)	
IV	13,193 (47.74)	16,579 (47.10)	18,451 (47.30)	20,105 (46.05)	22,457 (45.01)	
Unknown[5]	2,893 (10.47)	3,701 (10.51)	4,131 (10.59)	4,550 (10.42)	4,756 (9.53)	
Histology						<0.001
NSCLC NOS	10,701 (12.90)	14,483 (17.46)	16,652 (20.07)	18,940 (22.83)	22,177 (26.73)	
NSCLC Adenocarcinoma	11,470 (16.60)	13,331 (19.29)	13,880 (20.08)	14,828 (21.46)	15,600 (22.57)	
NSCLC Squamous	5,251 (12.64)	7,066 (17.01)	8,108 (19.52)	9,469 (22.80)	11,640 (28.03)	
NSCLC Large Cell	212 (11.84)	318 (17.76)	367 (20.49)	423 (23.62)	471 (26.30)	
Multiple Tumours						<0.001
1	21,867 (13.74)	28,293 (17.78)	31,702 (19.92)	35,939 (22.58)	41,329 (25.97)	
>1	5,767 (15.91)	6,905 (19.04)	7,305 (20.15)	7,721 (21.30)	8,559 (23.61)	
CCM (Between 78 to 6 Months Prior to Diagnosis)						<0.001
0	16,283 (58.92)	20,119 (57.16)	21,754 (55.77)	23,334 (53.44)	25,380 (50.87)	
1 - 2	8,003 (28.96)	10,533 (29.92)	11,911 (30.54)	13,907 (31.85)	16,670 (33.41)	
3+	3,348 (12.12)	4,546 (12.92)	5,342 (13.69)	6,419 (14.70)	7,838 (15.71)	

Appendix 4.6 Continued

	IMD 1[1] (Least Deprived) N (%)	IMD 2 N (%)	IMD 3 N (%)	IMD 4 N (%)	IMD 5 (Most Deprived) N (%)	P Value[2]
SACT Record						<0.001
Yes (In Time Range)[7]	6,719 (24.31)	7,972 (22.65)	8,457 (21.68)	8,967 (20.54)	9,883 (19.81)	
Yes (Not in Time Range)[7]	1,162 (4.2)	1,503 (4.3)	1,636 (4.2)	1,737 (4.0)	2,008 (4.0)	
No	19,753 (71.48)	25,723 (73.08)	28,914 (74.13)	32,956 (75.48)	37,997 (76.16)	
Treatment						
Utilised Chemotherapy	7,767 (28.11)	9,416 (26.75)	9,982 (25.59)	10,550 (24.16)	11,558 (23.17)	<0.001
Utilised Surgery	4,657 (16.85)	5,610 (15.94)	5,782 (14.82)	6,383 (14.62)	7,407 (14.85)	<0.001
Utilised Radiotherapy	7,647 (27.67)	9,524 (27.06)	10,757 (27.58)	11,910 (27.28)	13,997 (28.06)	0.014
Utilised Any Novel Anti-Cancer Therapy	1,847 (18.74)	2,029 (20.59)	2,015 (20.45)	1,992 (20.22)	1,971 (20.00)	<0.001
Utilised a Targeted Therapy	957 (20.01)	1,005 (21.02)	997 (20.84)	951 (19.88)	873 (18.25)	<0.001
Utilised a Biological	187 (18.00)	240 (23.10)	203 (19.54)	220 (21.17)	189 (18.19)	<0.001
Utilised an Immunotherapy	780 (17.74)	876 (19.92)	880 (20.01)	890 (20.24)	972 (22.10)	<0.001

[1]For diagnosis year 2012, IMD_2010 was used and for diagnosis years 2013 -2017, IMD_2015 was used.
[2]Chi-Square P Values
[3]Other refers to Asian/British Asian, Black/African/Caribbean/Black British, mixed/multiple ethnic groups and other ethnic group.
[4]Missing/unknown to unknown and missing ethnicity classifications.
[5]Unknown staging refers to missing and unstageable tumours
[6]Other refers to undifferentiated or anaplastic, undetermined, and missing tumour grades.
[7]In time range refers to an SACT record up to 56 days before and 2 years post diagnosis.
CCM: Charlson comorbidity; IMD: Index of Multiple Deprivation; NOS, Not Otherwise Specified; NSCLC, Non-Small Cell Lung Cancer; SACT; Systemic Anti-Cancer Therapy Dataset.

Appendix 4.7 Sensitivity analysis 1 for Chapter 4

Likelihood (OR and 95% CI and p values from logistic regression) of utilising any novel therapy (targeted therapy, immunotherapy, or biologic) by deprivation and adjusted for: sex, age, ethnicity, rural/urban indicator, stage, and comorbidities for patients with an adenocarcinoma histology NSCLC diagnosed between 01/01/2012 – 31/12/2017 (n = 69,109)

				Unadjusted			**Adjusted**		
	Number (%) Utilising a Novel Therapy n = 7,012 (10.15)	Number (%) Not Utilising a Novel Therapy n = 62,097 (89.85)	P Value[1]	OR	95% CI	P Value[2]	OR	95% CI	P Value[2]
IMD[3]			<0.001			**<0.001**			**<0.001**
1 (Least Deprived)	1,400 (12.21)	10,070 (87.79)		1.00	----- -----	-------	1.00	----- -----	-------
2	1,464 (10.98)	11,867 (89.02)		0.89	0.82 – 0.96	<0.001	0.86	0.80 – 0.94	<0.001
3	1,452 (10.46)	12,428 (89.54)		0.84	0.78 – 0.91	<0.001	0.78	0.72 – 0.85	<0.001
4	1,400 (9.44)	13,428 (90.56)		0.75	0.69 – 0.81	<0.001	0.67	0.61 – 0.72	<0.001
5 (Most Deprived)	1,296 (8.31)	14,304 (92.96)		0.65	0.60 – 0.71	<0.001	0.55	0.51 – 0.60	<0.001
Sex			<0.001						**<0.001**
Male	3,088 (8.99)	31,258 (91.01)		0.78	0.74 – 0.82	<0.001	0.74	0.70 – 0.77	<0.001
Female	3,924 (11.29)	30,839 (88.71)		1.00	----- -----	-------	1.00	----- -----	-------
Age (Years)			<0.001			**<0.001**			**<0.001**
<50	537 (20.61)	2,069 (79.39)		2.72	2.45 – 3.02	<0.001	2.09	1.88 – 2.34	<0.001
50 – 59	1,246 (14.38)	7,420 (85.62)		1.76	1.63 – 1.90	<0.001	1.50	1.39 – 1.63	<0.001
60 – 69	2,494 (11.44)	19,305 (88.56)		1.36	1.27 – 1.44	<0.001	1.27	1.19 – 1.35	<0.001
70 – 79	2,105 (8.70)	22,080 (91.30)		1.00	----- -----	-------	1.00	----- -----	-------
80 – 89	607 (5.57)	10,291 (99.43)		0.62	0.56 – 0.68	<0.001	0.61	0.55 – 0.67	<0.001
90+	23 (2.41)	932 (97.59)		0.26	0.17 – 0.39	<0.001	0.21	0.14 – 0.32	<0.001
Diagnosis Year			<0.001			**<0.001**			-------
2012	582 (5.54)	9,916 (94.46)		0.26	0.24 – 0.29	<0.001	-----	----- -----	-------
2013	740 (6.86)	10,042 (93.14)		0.33	0.30 – 0.36	<0.001	-----	----- -----	-------
2014	855 (7.50)	10,551 (92.50)		0.36	0.33 – 0.39	<0.001	-----	----- -----	-------
2015	1,054 (8.88)	10,814 (91.12)		0.43	0.40 – 0.47	<0.001	-----	----- -----	-------
2016	1,503 (12.31)	10,703 (87.69)		0.62	0.58 – 0.67	<0.001	-----	----- -----	-------
2017	2,278 (18.45)	10,071 (81.55)		1.00	----- -----	-------	-----	----- -----	-------

Appendix 4.7 Continued

	Number (%) Utilising a Novel Therapy n = 7,012 (10.15)	Number (%) Not Utilising a Novel Therapy n = 62,097 (89.85)	P Value[1]	Unadjusted OR	Unadjusted 95% CI	Unadjusted P Value[2]	Adjusted OR	Adjusted 95% CI	Adjusted P Value[2]
Ethnicity			<0.001			**<0.001**			**<0.001**
White	6,190 (9.73)	57,415 (90.27)		1.00	----- -----	-------	1.00	----- -----	-------
Other Ethnic Group[4]	667 (19.21)	2,806 (80.79)		2.20	2.02 – 2.41	<0.001	2.12	1.92 – 2.33	<0.001
Missing/unknown[5]	155 (7.63)	1,876 (92.37)		0.77	0.65 – 0.90	0.002	0.60	0.51 – 0.71	<0.001
Rural/Urban Indicator			0.046			**0.047**			**0.033**
Rural Village, Hamlet & Isolated Dwellings	645 (10.81)	5,319 (89.19)		1.06	0.97 – 1.16	0.216	0.91	0.82 – 1.00	0.051
Rural Town & Fringe	701 (10.53)	5,959 (89.47)		1.03	0.94 – 1.12	0.546	0.96	0.87 – 1.05	0.337
Urban City & Town	2,962 (9.82)	27,203 (90.18)		0.95	0.90 – 1.00	0.073	0.92	0.87 – 0.98	0.006
Urban Conurbation	2.704 (10.27)	23,615 (89.73)		1.00	----- -----	-------	1.00	----- -----	-------
Stage			<0.001			**<0.001**			**<0.001**
I	271 (1.99)	13,375 (98.01)		0.13	0.11 – 0.14	<0.001	0.13	0.12 – 0.15	<0.001
II	271 (5.27)	4,870 (94.73)		0.35	0.31 – 0.40	<0.001	0.36	0.32 – 0.41	<0.001
III	1,329 (12.30)	9,473 (87.70)		0.88	0.83 – 0.94	<0.001	0.88	0.83 – 0.94	<0.001
IV	5,014 (13.72)	31,527 (86.28)		1.00	----- -----	-------	1.00	----- -----	-------
Unknown[6]	127 (4.26)	2,852 (95.74)		0.28	0.23 – 0.34	<0.001	0.31	0.25 – 0.37	0.000
CCM (Between 78 to 6 Months Prior to Diagnosis)			<0.001			**<0.001**			**<0.001**
0	5,320 (12.13)	38,534 (87.87)		1.00	----- -----	-------	1.00	----- -----	-------
1-2	1,396 (7.43)	17,395 (92.57)		0.58	0.55 – 0.62	<0.001	0.70	0.70 – 0.80	<0.001
3+	296 (4.58)	6,168 (95.42)		0.35	0.31 – 0.39	<0.001	0.46	0.46 – 0.59	<0.001
Big Tumour Count			<0.001			**<0.001**			-------
1	5,998 (10.75)	49,773 (89.25)		1.00	----- -----	-------	-----	----- -----	-------
>1	1,014 (7.60)	12,324 (92.40)		0.68	0.64 – 0.73	<0.001	-----	----- -----	-------

Appendix 4.7 Continued

			Unadjusted				Adjusted				
	Number (%) Utilising a Novel Therapy n = 7,012 (10.15)	Number (%) Not Utilising a Novel Therapy n = 62,097 (89.85)	P Value[1]	OR	95% CI		P Value[2]	OR	95% CI		P Value[2]
Histology			-------				-------				-------
NSCLC NOS	----------	----------	-------	-----	-----	-----	-------	-----	-----	-----	-------
NSCLC Adenocarcinoma	7,012 (10.15)	62,097 (89.85)	-------	-----	-----	-----	-------	-----	-----	-----	-------
NSCLC Squamous	----------	----------	-------	-----	-----	-----	-------	-----	-----	-----	-------
NSCLC Large Cell	----------	----------	-------	-----	-----	-----	-------	-----	-----	-----	-------

[1]Chi-square P value

[2]P Values in bold are from likelihood ratio tests of the variable's contribution to the model. Unbolded P values are from a test of whether the OR is different from 1.

[3]For diagnosis year 2012, IMD_2010 was used and for diagnosis years 2013 - 2017, IMD_2015 was used.

[4]Other refers to Asian/British Asian, Black/African/Caribbean/Black British, mixed/multiple ethnic groups and other ethnic group.

[5]Missing/unknown ethnicity refers to missing and unknown ethnicity classifications.

[6]Unknown staging refers to missing and unstageable tumours.

[7]Other refers to undifferentiated or anaplastic, undetermined, and missing tumour grades.

CCM, Charlson comorbidity index; IMD, Index of Multiple Deprivation; NOS, Not otherwise specified; NSCLC; Non-Small Cell Lung Cancer; OR, Odds ratio; 95% CI, 95% Confidence interval.

Diagnosis year was not included in the model to improve fit.

Appendix 4.8 Sensitivity analysis 2 for Chapter 4

Likelihood (OR and 95% CI and p values from logistic regression) of utilising any novel anti-cancer therapy by deprivation and adjusted for: sex, age, ethnicity, rural/urban indicator, stage, and comorbidities for patients with a non-squamous NSCLC diagnosed between 01/01/2012 – 31/12/2017 (n = 153,853)

				Unadjusted			**Adjusted**		
	Number (%) Utilising Any Novel Anti-Cancer Therapy n = 8,123 (5.28)	Number (%) Not Utilising Any Novel Anti-Cancer Therapy n = 145,730 (94.72)	P Value[1]	OR	95% CI	P Value[2]	OR	95% CI	P Value[2]
IMD[3]			<0.001			**<0.001**			**<0.001**
1 (Least Deprived)	1,571 (7.02)	20,812 (92.98)		1.00	----- -----	-------	1.00	----- -----	-------
2	1,690 (6.01)	26,442 (93.99)		0.85	0.79 – 0.91	<0.001	0.83	0.77 – 0.89	<0.001
3	1,691 (5.47)	29,208 (94.53)		0.77	0.71 – 0.82	<0.001	0.72	0.67 – 0.78	<0.001
4	1,625 (4.75)	32,566 (95.25)		0.66	0.62 – 0.71	<0.001	0.59	0.54 – 0.63	<0.001
5 (Most Deprived)	1, 546 (4.04)	36,702 (95.96)		0.56	0.52 – 0.60	<0.001	0.46	0.43 – 0.50	<0.001
Sex			<0.001			**<0.001**			**<0.001**
Male	3,656 (4.65)	74,908 (95.35)		0.77	0.74 – 0.81	<0.001	0.73	0.70 – 0.77	<0.001
Female	4,467 (5.93)	70,822 (94.07)		1.00	----- -----	-------	1.00	----- -----	-------
Age (Years)			<0.001			**<0.001**			**<0.001**
<50	609 (16.91)	2,992 (83.09)		4.03	3.66 – 4.44	<0.001	2.90	2.62 – 3.20	<0.001
50 – 59	1,441 (11.11)	11,532 (88.89)		2.48	2.31 – 2.65	<0.001	1.99	1.86 – 2.14	<0.001
60 – 69	2,925 (8.03)	33,487 (91.97)		1.74	1.64 – 1.83	<0.001	1.56	1.47 – 1.65	<0.001
70 – 79	2,454 (4.80)	48,627 (95.20)		1.00	----- -----	-------	1.00	----- -----	-------
80 – 89	667 (1.62)	40.395 (98.38)		0.33	0.30 – 0.36	<0.001	0.35	0.32 – 0.39	<0.001
90+	27 (0.31)	8,697 (99.69)		0.06	0.04 – 0.09	<0.001	0.07	0.05 – 0.10	<0.001
Diagnosis Year			<0.001			**<0.001**			-------
2012	710 (2.85)	24,230 (97.15)		0.26	0.24 – 0.28	<0.001	-----	----- -----	-------
2013	868 (3.42)	24,541 (96.58)		0.31	0.29 – 0.34	<0.001	-----	----- -----	-------
2014	986 (3.83)	24,740 (96.17)		0.35	0.33 – 0.38	<0.001	-----	----- -----	-------
2015	1,215 (4.72)	24,511 (95.28)		0.44	0.41 – 0.47	<0.001	-----	----- -----	-------
2016	1,693 (6.50)	24,364 (93.50)		0.61	0.57 – 0.65	<0.001	-----	----- -----	-------
2017	2.651 (10.20)	23,344 (89.80)		1.00	----- -----	-------	-----	----- -----	-------

Appendix 4.8 Continued

	Number (%) Utilising Any Novel Anti-Cancer Therapy n = 8,123 (5.28)	Number (%) Not Utilising Any Novel Anti-Cancer Therapy n = 145,730 (94.72)	P Value[1]	Unadjusted OR	Unadjusted 95% CI	Unadjusted P Value[2]	Adjusted OR	Adjusted 95% CI	Adjusted P Value[2]
Ethnicity			<0.001			**<0.001**			**<0.001**
White	7,188 (5.12)	133,297 (94.88)		1.00	----- -----	-------	1.00	----- -----	-------
Other Ethnic Group[4]	754 (13.15)	4,978 (86.85)		2.81	2.59 – 3.04	<0.001	2.36	2.16 – 2.57	<0.001
Missing/unknown[5]	181 (2.37)	7,455 (97.63)		0.45	0.39 – 0.52	<0.001	0.47	0.40 – 0.54	<0.001
Rural/Urban Indicator			<0.001			**<0.001**			**<0.001**
Rural Village, Hamlet & Isolated Dwellings	755 (6.20)	11,423 (93.80)		1.14	1.05 – 1.24	0.001	0.91	0.84 – 1.00	0.049
Rural Town & Fringe	790 (5.31)	14,096 (94.69)		0.97	0.89 – 1.05	0.448	0.90	0.82 – 0.98	0.011
Urban City & Town	3.402 (4.95)	65,270 (95.05)		0.90	0.86 – 0.95	<0.001	0.87	0.82 – 0.92	<0.001
Urban Conurbation	3,176 (5.46)	54,941 (94.54)		1.00	----- -----	-------	1.00	----- -----	-------
Stage			<0.001			**<0.001**			**<0.001**
I	313 (1.24)	24,962 (98.76)		0.15	0.14 – 0.17	<0.001	0.17	0.15 – 0.19	<0.001
II	304 (3.07)	9,596 (96.93)		0.39	0.35 – 0.44	<0.001	0.43	0.38 – 0.48	<0.001
III	1,594 (6.75)	22,005 (93.25)		0.89	0.84 – 0.94	<0.001	0.93	0.88 – 0.99	0.019
IV	5,758 (7.51)	70,907 (92.49)		1.00	----- -----	-------	1.00	----- -----	-------
Unknown[6]	154 (0.84)	18,260 (99.16)		0.10	0.09 – 0.12	<0.001	0.16	0.14 – 0.19	<0.001
CCM (Between 78 to 6 Months Prior to Diagnosis)			<0.001			**<0.001**			**<0.001**
0	6,875 (6.59)	97,462 (93.41)		1.00	----- -----	-------	1.00	----- -----	-------
1-2	1,038 (2.96)	34,047 (97.04)		0.43	0.40 – 0.46	<0.001	0.65	0.61 – 0.68	<0.001
3+	210 (1.46)	14.221 (98.54)		0.21	0.18 – 0.24	<0.001	0.37	0.33 – 0.41	<0.001
Big Tumour Count			<0.001			**<0.001**			-------
1	6,915 (5.49)	119,083 (94.51)		1.00	----- -----	-------	-----	----- -----	-------
>1	1,208 (4.34)	26,647 (95.66)		0.78	0.73 – 0.83	<0.001	-----	----- -----	-------

Appendix 4.8 Continued

				Unadjusted			Adjusted			
	Number (%) Utilising Any Novel Anti-Cancer Therapy n = 8,123 (5.28)	Number (%) Not Utilising Any Novel Anti-Cancer Therapy n = 145,730 (94.72)	P Value[1]	OR	95% CI	P Value[2]	OR	95% CI		P Value[2]
Histology			<0.001			**<0.001**				-------
NSCLC NOS	1,082 (1.30)	81,871 (98.70)	-------	0.12	0.11 – 0.12	<0.001	-----	-----	-----	-------
NSCLC Adenocarcinoma	7,012 (10.15)	62,097 (89.85)	-------	1.00	----- -----	-------	-----	-----	-----	-------
NSCLC Squamous	----------	----------	-------	-----	----- -----	-------	-----	-----	-----	-------
NSCLC Large Cell	29 (1.62)	1,762 (98.38)	-------	0.15	0.10 – 0.21	<0.001	-----	-----	-----	-------

[1]Chi-square P value

[2]P Values in bold are from likelihood ratio tests of the variable's contribution to the model. Unbolded P values are from a test of whether the OR is different from 1.

[3]For diagnosis year 2012, IMD_2010 was used and for diagnosis years 2013 - 2017, IMD_2015 was used.

[4]Other refers to Asian/British Asian, Black/African/Caribbean/Black British, mixed/multiple ethnic groups and other ethnic groups.

[5]Missing/unknown refers to missing and unknown ethnicity classifications.

[6]Unknown staging refers to missing and unstageable tumours.

[7]Other refers to undifferentiated or anaplastic, undetermined, and missing tumour grades.

CCM, Charlson comorbidity index; IMD: Index of Multiple Deprivation; NOS: Not otherwise specified; NSCLC; Non-Small Cell Lung Cancer OR, Odds ratio; 95% CI, 95% Confidence interval.

Diagnosis year was dropped from the model to improve goodness of fit.

Appendix 4.9 Sensitivity analysis 3 for Chapter 4

Likelihood (OR and 95% CI and p values from logistic regression) of utilising any novel therapy (targeted therapy, immunotherapy, or biologic) by deprivation and adjusted for: age, diagnosis year, ethnicity, rural/urban indicator, stage, big tumour count, and comorbidities for patients with a NSCLC diagnosed between 01/04/2014 – 31/12/2017 (n = 122,708)

				Unadjusted			**Adjusted**		
	Number (%) Utilising a Novel Therapy n = 7,717 (6.29)	Number (%) Not Utilising a Novel Therapy n = 114,991 (93.71)	P Value[1]	OR	95% CI	P Value[2]	OR	95% CI	P Value[2]
IMD[3]			<0.001			**<0.001**			**<0.001**
1 (Least Deprived)	1,449 (8.21)	16,201 (91.79)		1.00	----- -----	-------	1.00	----- -----	-------
2	1,598 (7.24)	20,465 (92.76)		0.87	0.81 – 0.94	<0.001	0.88	0.81 – 0.95	0.002
3	1,579 (6.47)	22,814 (93.53)		0.77	0.72 – 0.83	<0.001	0.77	0.71 – 0.83	<0.001
4	1,558 (5.72)	25,668 (94.28)		0.68	0.63 – 0.73	<0.001	0.65	0.60 – 0.71	<0.001
5 (Most Deprived)	1,533 (4.89)	29.843 (95.11)		0.57	0.53 – 0.62	<0.001	0.53	0.49 – 0.58	<0.001
Sex			<0.001			**<0.001**			**<0.001**
Male	3,830 (5.80)	62,148 (94.20)		0.84	0.80 – 0.88	<0.001	0.83	0.79 – 0.88	0.000
Female	3,887 (6.85)	52,843 (93.15)		1.00	----- -----	-------	1.00	----- -----	-------
Age (Years)			<0.001			**<0.001**			**<0.001**
<50	513 (19.91)	2,064 (80.90)		4.17	3.75 – 4.63	<0.001	2.72	2.43 – 3.05	<0.001
50 – 59	1,301 (12.68)	8,960 (87.32)		2.43	2.27 – 2.61	<0.001	1.75	1.62 – 1.89	<0.001
60 – 69	2,796 (9.30)	27,254 (90.70)		1.72	1.63 – 1.82	<0.001	1.47	1.38 – 1.56	<0.001
70 – 79	2,464 (5.63)	41,314 (94.37)		1.00	----- -----	-------	1.00	----- -----	-------
80 – 89	620 (2.05)	29,609 (97.95)		0.35	0.32 – 0.38	<0.001	0.47	0.43 – 0.51	<0.001
90+	23 (0.40)	5,790 (99.60)		0.07	0.04 – 0.10	<0.001	0.12	0.08 – 0.19	<0.001
Diagnosis Year			<0.001			**<0.001**			**<0.001**
2012	----------	----------		------	----- -----	-------	-----	----- -----	-------
2013	----------	----------		-----	----- -----	-------	-----	----- -----	-------
2014	844 (3.45)	23,596 (96.55)		0.31	0.29 – 0.34	<0.001	0.28	0.26 – 0.30	<0.001
2015	1,439 (4.40)	31,230 (95.60)		0.40	0.38 – 0.43	<0.001	0.36	0.33 – 0.38	<0.001
2016	2,097 (6.36))	30,889 (93.64)		0.60	0.56 – 0.63	<0.001	0.55	0.52 – 0.58	<0.001
2017	3,337 (10.23)	29,276 (89.77)		1.00	----- -----	-------	1.00	----- -----	-------

Appendix 4.9 Continued

	Number (%) Utilising a Novel Therapy n = 7,717 (6.29)	Number (%) Not Utilising a Novel Therapy n = 114,991 (93.71)	P Value[1]	Unadjusted OR	Unadjusted 95% CI	Unadjusted P Value[2]	Adjusted OR	Adjusted 95% CI	Adjusted P Value[2]
Ethnicity			<0.001			**<0.001**			**<0.001**
White	6,903 (6.17)	105,024 (93.83)		1.00	----- -----	-------	1.00	----- -----	-------
Other Ethnic Group[4]	618 (13.62)	3,921 (86.38)		2.40	2.20 – 2.62	<0.001	1.90	1.72 – 2.10	<0.001
Missing/unknown[5]	196 (3.14)	6,046 (96.86)		0.49	0.43 – 0.57	<0.001	0.55	0.47 – 0.64	<0.001
Rural/Urban Indicator			<0.001			**<0.001**			**<0.001**
Rural Village, Hamlet & Isolated Dwellings	739 (7.52)	9,089 (92.48)		1.21	1.11 – 1.31	<0.001	0.94	0.85 – 1.03	0.183
Rural Town & Fringe	753 (6.29)	11,220 (93.71)		1.00	0.92 – 1.08	0.933	0.90	0.82 – 0.98	0.020
Urban City & Town	3,299 (6.05)	51,237 (93.95)		0.96	0.91 – 1.01	0.086	0.92	0.87 – 0.97	0.003
Urban Conurbation	2,926 (6.31)	43.445 (93.69)		1.00	----- -----	-------	1.00	----- -----	-------
Stage			<0.001			**<0.001**			**<0.001**
I	313 (1.41)	21,960 (98.59)		0.15	0.13 – 0.16	<0.001	0.13	0.11 – 0.15	<0.001
II	348 (3.56)	9.429 (96.44)		0.38	0.34 – 0.42	<0.001	0.38	0.34 – 0.43	<0.001
III	1,864 (8.03)	21,337 (91.97)		0.90	0.85 – 0.95	<0.001	0.96	0.90 – 1.01	0.131
IV	5,089 (8.88)	52,189 (91.12)		1.00	----- -----	-------	1.00	----- -----	-------
Unknown[6]	103 (1.01)	10.076 (98.99)		0.10	0.09 – 0.13	<0.001	0.26	0.21 – 0.31	<0.001
CCM (Between 78 to 6 Months Prior to Diagnosis)			<0.001			**<0.001**			**<0.001**
0	5,610 (8.63)	59,375 (91.37)		1.00	----- -----	-------	1.00	----- -----	-------
1-2	1,712 (4.39)	37,301 (95.61)		0.49	0.46 – 0.51	<0.001	0.71	0.67 – 0.76	<0.001
3+	395 (2.11)	18,315 (97.89)		0.23	0.21 – 0.25	<0.001	0.45	0.41 – 0.51	<0.001
Big Tumour Count			<0.001			**<0.001**			-------
1	6,506 (6.54)	92,953 (93.46)		1.00	----- -----	-------	-----	----- -----	-------
>1	1,211 (5.21)	22,038 (94.79)		0.79	0.74 – 0.84	<0.001	-----	----- -----	-------

Appendix 4.9 Continued

				Unadjusted			Adjusted		
	Number (%) Utilising a Novel Therapy n = 7,717 (6.29)	Number (%) Not Utilising a Novel Therapy n = 114,991 (93.71)	P Value[1]	OR	95% CI	P Value[2]	OR	95% CI	P Value[2]
Histology			<0.001			**<0.001**			**<0.001**
NSCLC NOS	800 (1.57)	50,084 (98.43)		0.12	0.11 – 0.12	-------	0.18	0.16 – 0.19	<0.001
NSCLC Adenocarcinoma	5,482 (12.17)	39,560 (87.83)		1.00	----- -----	-------	1.00	----- -----	-------
NSCLC Squamous	1,417 (5.52)	24,247 (94.48)		0.42	0.40 – 0.45	-------	0.50	0.46 – 0.53	<0.001
NSCLC Large Cell	18 (1.61)	1,100 (98.39		0.12	0.07 – 0.19	-------	0.10	0.06 – 0.17	<0.001

[1]Chi-square P value

[2]P Values in bold are from likelihood ratio tests of the variable's contribution to the model. Unbolded P values are from a test of whether the OR is different from 1.

[3]For diagnosis year 2010, IMD_2010 was used and for diagnosis years 2013 -2017, IMD_2015 was used.

[4]Other refers to Asian/British Asian, Black/African/Caribbean/Black British, mixed/multiple ethnic groups and other ethnic groups.

[5]Missing/unknown ethnicity refers to missing and unknown ethnicity classifications.

[6]Unknown staging refers to missing and unstageable tumours.

[7]Other refers to undifferentiated or anaplastic, undetermined, and missing tumour grades.

CCM, Charlson comorbidity index; ER, Oestrogen receptor status; IMD: Index of Multiple Deprivation; NOS: Not Otherwise Specified; NSCLC; Non-Small Cell Lung Cancer OR, Odds ratio; 95% CI, 95% Confidence interval.

Appendix 4.10 Clinical and demographic descriptive by treatment utilisation for Chapter 4

	All Patients n = 195,387 (100%)	All Patients (%) Utilising a Novel Therapy n = 9,854 (5.04)	All Patients (%) Utilising Any Targeted Therapy n = 4,783 (2.45)	All Patients (%) Utilising An EGFRi n = 4,234 (2.17)	All Patients (%) Utilising an ALKi n = 540 (0.28)	All Patients (%) Utilising a Biological n = 1,039 (0.53)	All Patients (%) Utilising an Immunotherapy n = 4,398 (2.25)
IMD[1]							
1 (Least Deprived)	27,634 (14.14)	1,847 (6.68)	957 (3.46)	840 (3.04)	117 (0.42)	187 (0.68)	780 (2.82)
2	35,198 (18.01)	2,029 (5.76)	1,005 (2.86)	879 (2.50)	125 (0.36)	240 (0.68)	876 (2.49)
3	39,007 (19.96)	2,015 (5.17)	997 (2.56)	884 (2.27)	107 (0.27)	203 (0.52)	880 (2.26)
4	43,660 (22.35)	1,992 (4.56)	951 (2.18)	845 (1.94)	103 (0.24)	220 (0.50)	890 (2.04)
5 (Most Deprived)	49,888 (25.53)	1,971 (3.95)	873 (1.75)	786 (1.58)	88 (0.18)	189 (0.38)	972 (1.95)
Sex							
Male	105,717 (54.11)	4,835 (4.57)	2,016 (1.91)	1,753 (1.66)	257 (0.24)	541 (0.51)	2,441 (2.31)
Female	89,670 (45.89)	5,019 (5.60)	2,767 (3.09)	2,481 (2.77)	283 (0.32)	498 (0.56)	1,957 (2.18)
Age (Years)							
<50	4,206 (2.15)	673 (16.00)	406 (9.65)	265 (6.30)	142 (3.38)	93 (2.21)	216 (5.14)
50 – 59	16,609 (8.50)	1,686 (10.15)	776 (4.67)	657 (3.96)	117 (0.70)	246 (0.88)	746 (4.49)
60 – 69	48,561 (24.85)	3,602 (7.42)	1,648 (3.39)	1,503 (3.10)	142 (0.29)	429 (0.88)	1,674 (3.45)
70 – 79	68,294 (34.95)	3,080 (4.51)	1,421 (2.08)	1,310 (1.92)	107 (0.16)	248 (0.36)	1,497 (2.19)
80 – 89	48,557 (24.85)	784 (1.61)	505 (1.04)	472 (0.97)	32 (0.07)	23 (0.05)	263 (0.54)
90+	9,160 (4.69)	29 (0.32)	27 (0.29)	27 (0.29)	0 (0.00)	0 (0.00)	2 (0.02)
Diagnosis Year							
2012	32,130 (16.44)	861 (2.68)	847 (2.64)	832 (2.59)	15 (0.05)	8 (0.02)	4 (0.01)
2013	32,428 (16.60)	991 (3.06)	930 (2.87)	891 (2.75)	40 (0.12)	25 (0.08)	34 (0.10)
2014	32,561 (16.66)	1,129 (3.47)	917 (2.82)	848 (2.60)	69 (0.21)	118 (0.36)	108 (0.33)
2015	32,669 (16.72)	1,439 (4.40)	744 (2.28)	643 (1.97)	99 (0.30)	328 (1.00)	422 (1.29)
2016	32,986 (16.88)	2,097 (6.36)	598 (1.81)	459 (1.39)	135 (0.41)	302 (0.92)	1,315 (3.99)
2017	32,613 (16.69)	3,337 (10.23)	747 (2.29)	561 (1.72)	182 (0.56)	258 (0.79)	2,515 (7.71)

Appendix 4.10 Continued

	All Patients n = 195,387 (100%)	All Patients (%) Utilising a Novel Therapy n = 9,854 (5.04)	All Patients (%) Utilising Any Targeted Therapy n = 4,783 (2.45)	All Patients (%) Utilising An EGFRi n = 4,234 (2.17)	All Patients (%) Utilising an ALKi n = 540 (0.28)	All Patients (%) Utilising a Biological n = 1,039 (0.53)	All Patients (%) Utilising an Immunotherapy n = 4,398 (2.25)
Ethnicity							
White	179,690 (91.97)	8,812 (4.90)	4,142 (2.31)	3,704 (2.06)	433 (0.24)	940 (0.52)	4,044 (2.25)
Other Ethnic Group[2]	7,007 (3.59)	821 (11.72)	551 (7.86)	453 (6.46)	94 (1.34)	80 (1.14)	229 (3.27)
Missing/unknown[3]	8,690 (4.45)	221 (2.54)	90 (1.04)	453 (6.46)	13 (0.15)	19 (0.22)	125 (1.44)
Rural/Urban Indicator							
Rural Village, Hamlet & Isolated Dwellings	15,384 (7.87)	911 (5.92)	427 (2.78)	383 (2.49)	44 (0.29)	112 (0.73)	417 (2.71)
Rural Town & Fringe	18,798 (9.62)	950 (5.05)	460 (2.45)	401 (2.13)	57 (0.30)	103 (0.55)	424 (2.26)
Urban City & Town	86,505 (44.27)	4,126 (4.77)	1,885 (2.18)	1,678 (1.94)	205 (0.24)	441 (0.51)	1,931 (2.23)
Urban Conurbation	74,700 (38.23)	3,867 (5.18)	2,011 (2.69)	1,772 (2.37)	234 (0.31)	383 (0.51)	1,626 (2.18)
Stage							
I	32,408 (16.59)	384 (1.18)	185 (0.57)	173 (0.53)	8 (0.02)	33 (0.10)	177 (0.55)
II	15,321 (7.84)	437 (2.85)	167 (1.09)	152 (0.99)	14 (0.09)	32 (0.21)	247 (1.61)
III	36,842 (18.86)	2,307 (6.26)	832 (2.26)	738 (2.00)	94 (0.26)	204 (0.55)	1,324 (3.59)
IV	90,785 (46.46)	6,545 (7.21)	3,484 (3.84)	3,059 (3.37)	421 (0.46)	746 (0.82)	2,604 (2.87)
Unknown[4]	20,031 (10.25)	181 (0.90)	115 (0.57)	112 (0.56)	3 (0.01)	24 (0.12)	46 (0.23)
Multiple Tumours							
1	159,130 (81.44)	8,340 (5.24)	4,096 (2.57)	3,611 (2.27)	483 (0.30)	867 (0.54)	3,682 (2.31)
>1	36,257 (18.56)	1,514 (4.18)	687 (1.89)	623 (1.72)	57 (0.16)	172 (0.47)	716 (1.97)
CCM (Between 78 to 6 Months Prior to Diagnosis)							
0	106,870 (54.70)	7,248 (6.78)	3,675 (3.44)	3,225 (3.02)	445 (0.42)	779 (0.73)	3,078 (2.88)
1-2	61,024 (31.23)	2,134 (3.50)	917 (1.50)	827 (1.36)	88 (0.14)	220 (0.36)	1,072 (1.76)
3+	27,493 (14.07)	472 (1.72)	191 (0.69)	182 (0.66)	7 (0.03)	40 (0.15)	248 (0.90)

Appendix 4.10 Continued

	All Patients n = 195,387 (100%)	All Patients (%) Utilising a Novel Therapy n = 9,854 (5.04)	All Patients (%) Utilising Any Targeted Therapy n = 4,783 (2.45)	All Patients (%) Utilising An EGFRi n = 4,234 (2.17)	All Patients (%) Utilising an ALKi n = 540 (0.28)	All Patients (%) Utilising a Biological n = 1,039 (0.53)	All Patients (%) Utilising an Immunotherapy n = 4,398 (2.25)
Histology							
NSCLC (Including 8000 & 8010 Code)	82,953 (42.46)	1,082 (1.30)	539 (0.65)	481 (0.58)	56 (0.07)	111 (0.13)	474 (0.57)
NSCLC Adenocarcinoma	69,109 (35.37)	7,012 (10.15)	3,835 (5.55)	3,360 (4.86)	469 (0.68)	882 (1.28)	2,610 (3.78)
NSCLC Squamous	41,534 (21.26)	1,731 (4.17)	394 (0.95)	380 (0.91)	13 (0.03)	39 (0.09)	1,304 (3.14)
NSCLC Large Cell	1,791 (0.92)	29 (1.62)	15 (0.84)	13 (0.73)	2 (0.11)	7 (0.39)	10 (0.56)
Other Treatments							
Utilised Chemotherapy	49.273 (25.22)	9,155 (18.58)	4,497 (94.02)	3,967 (93.69)	526 (97.41)	1,003 (96.54)	4,011 (91.20)
Utilised Surgery	29.839 (15.27)	932 (9.46)	395 (8.26)	363 (8.57)	30 (5.56)	86 (8.28)	482 (10.96)
Utilised Radiotherapy	53,835 (27.55)	3,362 (34.13)	1,550 (32.41)	1,399 (33.04)	147 (27.22)	307 (29.55)	1,582 (35.97)
Utilised Chemotherapy & Surgery	7,345 (3.76)	595 (6.04)	252 (5.27)	227 (5.36)	25 (4.63)	67 (6.45)	299 (6.80)
Utilise Chemotherapy & Radiotherapy	19,695 (10.08)	3,092 (31.38)	1,455 (30.42)	1,311 (30.96)	141 (26.11)	296 (28.49)	1,415 (32.17)
Utilised Surgery & Radiotherapy	1,964 (1.01)	124 (1.26)	45 (0.94)	40 (0.94)	5 (0.93)	12 (1.15)	73 (1.66)
Utilised All Conventional Therapies	783 (0.40)	91 (0.40)	32 (0.67)	27 (0.64)	5 (0.93)	9 (0.87)	54 (1.23)
Utilised Any Conventional Therapy	104,726 (53.60)	9,730 (98.74)	4,722 (98.72)	4,178 (98.68)	537 (99.44)	1,030 (99.13)	4,342 (98.73)

[1]For diagnosis year 2012, IMD_2010 was used and for diagnosis years 2013 -2017, IMD_2015 was used.
[2]Other refers to Asian/British Asian, Black/African/Caribbean/Black British, mixed/multiple ethnic groups and other ethnic groups.
[3]Missing/unknown ethnicity refers to unknown and missing ethnicity classifications.
[4]Unknown staging refers to missing and unstageable tumours
ALKi: Anaplastic lymphoma kinase inhibitor; CCM: Charlson Comorbidity Index; EGFRi: Epidermal growth factor receptor inhibitor; IMD: Index of multiple deprivation; n: Number; NSCLC: Non-small cell lung cancer.

Appendix 4.11 Any novel therapy sub-group analyses for Chapter 4

Likelihood (adjOR and 95% CI and p values from logistic regression) of utilising various novel anti-cancer therapy sub classifications by deprivation and adjusted for: sex, age (<50, 60-69, 70-79, 80+), and ethnicity in patients with a NSCLC diagnosed between 01/01/2012 – 31/12/2017 (n = 195,387)

	Any Novel Therapy n = 9,854 (5.04)			**Any Targeted Therapy (EGFRi, ALKi or Other)** n = 4,783 (2.45)			**Targeted Therapy (EGFRi)** n = 4,234 (2.17)		
	adjOR	95% CI	P Value[3]	adjOR	95% CI	P Value[3]	adjOR	95% CI	P Value[3]
IMD[2]			**<0.001**			**<0.001**			**<0.001**
1 (Least Deprived)	1.00	----- -----	-------	1.00	----- -----	-------	1.00	----- -----	-------
2	0.84	0.78 – 0.89	0.000	0.81	0.74 – 0.88	0.000	0.81	0.73 – 0.89	0.000
3	0.71	0.66 – 0.76	0.000	0.68	0.62 – 0.74	0.000	0.69	0.63 – 0.76	0.000
4	0.58	0.55 – 0.62	0.000	0.54	0.49 – 0.59	0.000	0.56	0.51 – 0.61	0.000
5 (Most Deprived)	0.46	0.43 – 0.49	0.000	0.40	0.36 – 0.44	0.000	0.42	0.38 – 0.46	0.000

	Targeted Therapy (EGFRi no Erlotinib) n = 1,827 (0.94)			**Targeted Therapy (ALKi)** n = 540 (0.28)			**Biologics[1]** n = 1,039 (0.53)		
	adjOR	95% CI	P Value[3]	adjOR	95% CI	P Value[3]	adjOR	95% CI	P Value[3]
IMD[2]			**<0.001**			**<0.001**			**<0.001**
1 (Least Deprived)	1.00	----- -----	-------	1.00	----- -----	-------	1.00	----- -----	-------
2	0.76	0.66 – 0.87	0.000	0.83	0.64 – 1.06	0.138	0.98	0.81 – 1.19	0.854
3	0.59	0.52 – 0.68	0.000	0.57	0.44 – 0.75	0.000	0.71	0.58 – 0.86	0.001
4	0.47	0.41 – 0.55	0.000	0.44	0.34 – 0.57	0.000	0.64	0.52 – 0.78	0.000
5 (Most Deprived)	0.32	0.27 – 0.37	0.000	0.29	0.22 – 0.38	0.000	0.43	0.35 – 0.53	0.000

Appendix 4.11 Continued

	Immunotherapy[1] n = 4,398 (2.25) adjOR	95% CI	P Value[3]
IMD[2]			**<0.001**
1 (Least Deprived)	1.00	----- -----	-------
2	0.86	0.78 – 0.95	0.003
3	0.75	0.68 – 0.83	0.000
4	0.65	0.59 – 0.72	0.000
5 (Most Deprived)	0.58	0.53 – 0.64	0.000

[1]Biologics and immunotherapy models are not adjusted for sex as this variable was not significant
[2]For diagnosis year 2012, IMD_2010 was used and for diagnosis years 2013 - 2017, IMD_2015 was used.
[3]Bolded P values are from likelihood ratio tests of the variable's contribution to the model. Unbolded P values are from a test of whether the adjOR is different from 1.
adjOR: Adjusted odds ratio: ALKi: Anaplastic lymphoma kinase inhibitors; EGFRi: Epidermal growth factor receptor inhibitors; IMD: Index of multiple deprivation; n: Number; NSCLC: Non-small cell lung cancer; 95% CI: 95% Confidence interval.

References

1. Langreth R, Waldholz M. New era of personalized medicine: targeting drugs for each unique genetic profile. Oncologist. 1999;4(5):426–7.

2. Joyner MJ, Paneth N. Promises, promises, and precision medicine. J Clin Invest. 2019;129(3):946–8.

3. Jørgensen JT. Twenty years with personalized medicine: past, present, and future of individualized pharmacotherapy. Oncologist. 2019;24(7):e432–40.

4. Remon J, Dienstmann R. Precision oncology: separating the wheat from the chaff. ESMO Open. 2018;3(6):e000446.

5. Williams E, Buck D, Babalola G. The King's Fund: What are health inequalities? [Internet] 2022 [cited March 27th 2022]. Available from: https://www.kingsfund.org.uk/publications/what-are-health-inequalities.

6. Starfield B, Robertson J, Riley AW. Social class gradients and health in childhood. Ambul Pediatr Off J Ambul Pediatr Assoc. 2002;2(4):238–46.

7. Smith JP, Kington R. Demographic and economic correlates of health in old age. Demography. 1997;34(1):159–70.

8. Fiscella K. Is lower income associated with greater biopsychosocial morbidity? Implications for physicians working with underserved patients. J Fam Pract. 1999;48(5):372–7.

9. Wong MD, Shapiro MF, Boscardin WJ, Ettner SL. Contribution of major diseases to disparities in mortality. N Engl J Med. 2002;347(20):1585–92.

10. Marmot M. Introduction. In: Marmot M and Wilkinson R, editors. The social determinants of health. Oxford: Oxford University Press; 2006. p.1–5.

11. Murage P, Crawford SM, Bachmann M, Jones A. Geographical disparities in access to cancer management and treatment services in England. Health Place. 2016;42:11–8.

12. Marmot M. Fair society, healthy lives (The Marmot Review): strategic review of health inequalities in England post-2010 [Internet]. 2010 [cited October 7th 2020]. Available from: https://www.parliament.uk/globalassets/documents/fair-society-healthy-lives-full-report.pdf.

13. Popham F, Dibben C, Bambra C. Are health inequalities really not the smallest in the Nordic welfare states? A comparison of mortality inequality in 37 countries. J Epidemiol Community Health. 2013;67(5):412–8.

14. Mackenbach JP. The persistence of health inequalities in modern welfare states: the explanation of a paradox. Soc Sci Med. 2012;75(4):761–9.

15. Baker EH. Socioeconomic status, definition. In: Cockerham WC and Quah S, editors. Wiley Blackwell Encycl Health Illn Behav Soc. 2014. p. 2210–4.

16. Bambra C. Work, worklessness, and the political economy of health. Oxford: Oxford University Press; 2011.

17. Mueller CW, Parcel TL. Measures of socioeconomic status: alternatives and recommendations. Child Dev. 1981;52(1):13–30.

18. Shavers VL, Brown ML. Racial and ethnic disparities in the receipt of cancer treatment. J Natl Cancer Inst. 2002;94(5):334–57.

19. Galobardes B, Shaw M, Lawlor DA, Lynch JW, Davey Smith G. Indicators of socioeconomic position (part 1). J Epidemiol Community Health. 2006;60(1):7–12.

20. Krieger N, Williams DR, Moss NE. Measuring social class in US public health research: concepts, methodologies, and guidelines. Annu Rev Public Health. 1997;18:341–78.

21. Braveman PA, Cubbin C, Egerter S, Chideya S, Marchi KS, Metzler M, *et al*. Socioeconomic status in health research: one size does not fit all. JAMA. 2005;294(22):2879–88.

22. Shavers VL. Measurement of socioeconomic status in health disparities research. J Natl Med Assoc. 2007;99(9):1013–23.

23. Adler NE, Newman K. Socioeconomic disparities in health: pathways and policies. Health Aff Proj Hope. 2002;21(2):60–76.

24. Turrell G. Income non-reporting: implications for health inequalities research. J Epidemiol Community Health. 2000;54(3):207–14.

25. Liberatos P, Link BG, Kelsey JL. The measurement of social class in epidemiology. Epidemiol Rev. 1988;10:87–121.

26. Fiscella K, Williams DR. Health disparities based on socioeconomic inequities: implications for urban health care. Acad Med. 2004;79(12):1139–47.

27. Link BG, Phelan J. Social conditions as fundamental causes of disease. J Health Soc Behav. 1995;80–94.

28. Public Health England. Health profile for England: 2018 [Internet]. 2018 [cited March 31st 2022]. Available from: https://www.gov.uk/government/publications/health-profile-for-england-2018/chapter-5-inequalities-in-health#wider-determinants-of-health.

29. World Health Organisation. Determinants of health [Internet]. 2017 [cited March 31st 2022]. Available from: https://www.who.int/news-room/questions-and-answers/item/determinants-of-health.

30. Dahlgren G, Whitehead M. Tackling inequalities in health: what can we learn from what has been tried? Working paper prepared for the King's Fund international seminar on tackling inequalities in health. In: Dahlgren G and Whitehead M, editors. European strategies for tackling social inequities in health: levelling up part 2 Copenhagen: WHO Regional Office for Europe. 1993. p. 20–32.

31. Bambra C. Health divides where you live can kill you. Bristol: Policy Press; 2016.

32. Bourdieu P. The forms of capital. In Richardson JG, editor. Handbook of theory and research for sociology education. New York: Greenwood Press; 1986. p. 241–58.

33. Williams DR, Mohammed SA, Leavell J, Collins C. Race, socioeconomic status and health: complexities, ongoing challenges and research opportunities. Ann N Y Acad Sci. 2010;1186:69–101.

34. Acheson D. Independent inquiry into inequalities in health report [Internet]. 1998 [cited February 22nd 2020]. Available from: https://www.gov.uk/government/publications/independent-inquiry-into-inequalities-in-health-report.

35. Cerutti J, Lussier AA, Zhu Y, Liu J, Dunn EC. Associations between indicators of socioeconomic position and DNA methylation: a scoping review. Clin Epigenetics. 2021;13(1):221.

36. Mathers JC, Strathdee G, Relton CL. Induction of epigenetic alterations by dietary and other environmental factors. Adv Genet. 2010;71:3–39.

37. Goldberg DS. The implications of fundamental cause theory for priority setting. Am J Public Health. 2014;104(10):1839–43.

38. Bartley M. Health inequality: an introduction to concepts, theories and methods. Second Edition. Cambridge: Polity Press; 2017.

39. Wegener B. Job mobility and social ties: social resources, prior job, and status attainment. Am Soc Rev. 1960;56:60–71.

40. McEwen BS. Stress, adaptation, and disease. Allostasis and allostatic load. Ann N Y Acad Sci. 1998;840:33–44.

41. Lynch JW, Kaplan GA, Shema SJ. Cumulative impact of sustained economic hardship on physical, cognitive, psychological, and social functioning. N Engl J Med. 1997;337(26):1889–95.

42. Bosma H, van de Mheen HD, Mackenbach JP. Social class in childhood and general health in adulthood: questionnaire study of contribution of psychological attributes. BMJ. 1999;318(7175):18–22.

43. Adler NE, Page AE. The psychosocial needs of cancer patients. In: Nancy AE and Page AEK, editors. Cancer care for the whole patient: meeting psychosocial health needs. Washington (DC): National Academies Press (US); 2008.

44. Jones NL, Gilman SE, Cheng TL, Drury SS, Hill CV, Geronimus AT. Life course approaches to the causes of health disparities. Am J Public Health. 2019;109(Suppl 1):S48–55.

45. Kawachi I, Subramanian SV, Almeida-Filho N. A glossary for health inequalities. J Epidemiol Community Health. 2002;56(9):647–52.

46. Skalická V, van Lenthe F, Bambra C, Krokstad S, Mackenbach J. Material, psychosocial, behavioural and biomedical factors in the explanation of relative socio-economic inequalities in mortality: evidence from the HUNT study. Int J Epidemiol. 2009;38(5):1272–84.

47. van Oort FVA, van Lenthe FJ, Mackenbach JP. Material, psychosocial, and behavioural factors in the explanation of educational inequalities in mortality in the Netherlands. J Epidemiol Community Health. 2005;59(3):214–20.

48. Duijster D, Oude Groeniger J, van der Heijden GJMG, van Lenthe FJ. Material, behavioural, cultural and psychosocial factors in the explanation of socioeconomic inequalities in oral health. Eur J Public Health. 2018;28(4):590–7.

49. Arcaya MC, Arcaya AL, Subramanian SV. Inequalities in health: definitions, concepts, and theories. Glob Health Action. 2015;8:10.3402/gha.v8.27106.

50. Smith GD, Hart C, Watt G, Hole D, Hawthorne V. Individual social class, area-based deprivation, cardiovascular disease risk factors, and mortality: the Renfrew and Paisley Study. J Epidemiol Community Health. 1998;52(6):399.

51. McCartney G, Collins C, Mackenzie M. What (or who) causes health inequalities: theories, evidence and implications? Health Policy Amst Neth. 2013;113(3):221–7.

52. Madhusoodanan J. Health-care inequality could deepen with precision oncology. Nature. 2020;585(7826):S13–5.

53. Osarogiagbon RU, Sineshaw HM, Unger JM, Acuña-Villaorduña A, Goel S. Immune-based cancer treatment: addressing disparities in access and outcomes. Am Soc Clin Oncol Educ Book. 2021;(41):66–78.

54. Braveman P, Gottlieb L. The social determinants of health: it's time to consider the causes of the causes. Public Health Rep. 2014;129(Suppl 2):19–31.

55. Hart JT. The inverse care law. Lancet. 1971;1(7696):405–12.

56. White M, Adams J, Heywood P. How and why do interventions that increase health overall widen inequalities within populations? In: Babones SJ, editor. Social inequality and public health. Bristol: Policy Press; 2012.

57. James JE. Personalised medicine, disease prevention, and the inverse care law: more harm than benefit? Eur J Epidemiol. 2014;29(6):383–90.

58. Marmot M. Inequalities in health. N Engl J Med. 2001;345(2):134–6.

59. Evans GW. The environment of childhood poverty. Am Psychol. 2004;59(2):77–92.

60. Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, *et al*. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin. 2021;71(3):209–49.

61. Ahmad AS, Ormiston-Smith N, Sasieni PD. Trends in the lifetime risk of developing cancer in Great Britain: comparison of risk for those born from 1930 to 1960. Br J Cancer. 2015;112(5):943-7.

62. Su M, Zhang Z, Zhou L, Han C, Huang C, Nice EC. Proteomics, personalized medicine and cancer. Cancers. 2021;13(11):2512.

63. World Health Organisation. Cancer [Internet]. 2022 [cited January 26th 2022]. Available from: https://www.who.int/news-room/fact-sheets/detail/cancer.

64. World Health Organisation. International classification of disease for oncology third edition [Internet]. 2013 [cited February 16th 2022]. Available from: https://apps.who.int/iris/bitstream/handle/10665/96612/9789241548496_eng.pdf.

65. National Cancer Institute. Cancer classification: SEER training module [Internet]. 2022 [cited January 26th 2022]. Available from: https://training.seer.cancer.gov/disease/categories/classification.html.

66. National Cancer Institute. What is cancer? [Internet]. 2021 [cited February 16th 2022]. Available from: https://www.cancer.gov/about-cancer/understanding/what-is-cancer.

67. Chow A. Cell cycle control by oncogenes and tumor suppressors: driving the transformation of normal cells into cancerous cells. Nat Educ. 2010;3(9):7.

68. National Institute for Health and Care Excellence. Trastuzumab nation funding/access decisions [Internet]. 2022 [cited April 3rd 2022]. Available from: https://bnf.nice.org.uk/drug/trastuzumab.html#nationalFunding.

69. Kiernan U, Naylor S. Precision medicine drugs pleonasm or reality? Drug Discov World. 2018;9–16.

70. Mayor S. WHO includes 16 new cancer drugs on list of essential medicines. Lancet Oncol. 2015;16(7):P757.

71. Sundar R, Chénard-Poirier M, Collins DC, Yap TA. Imprecision in the era of precision medicine in non-small cell lung cancer. Front Med. 2017;4:39.

72. Herbst R, Morgensztern D, Boshoff C. The biology and management of non-small cell lung cancer. Nature. 2018;553:446–54.

73. Politi K, Herbst RS. Lung cancer in the era of precision medicine. Clin Cancer Res. 2015;21(10):2213–20.

74. McDonald ES, Clark AS, Tchou J, Zhang P, Freedman GM. Clinical diagnosis and management of breast cancer. J Nucl Med. 2016;57(Suppl) 1:9S-16S.

75. Harkness EF, Astley SM, Evans DG. Risk-based breast cancer screening strategies in women. Best Pract Res Clin Obstet Gynaecol. 2020;65:3–17.

76. NHS digital national statistics. NHS breast screening programme [Internet]. 2022 [cited Mar 29th 2022]. Available from: https://digital.nhs.uk/data-and-information/publications/statistical/breast-screening-programme/england---2020-21/mainreport6#section-1-coverage.

77. Dai X, Li T, Bai Z, Yang Y, Liu X, Zhan J, *et al.* Breast cancer intrinsic subtype classification, clinical use and future trends. Am J Cancer Res. 2015;5(10):2929–43.

78. Akinyemiju TF, Pisu M, Waterbor JW, Altekruse SF. Socioeconomic status and incidence of breast cancer by hormone receptor subtype. Springerplus. 2015;4:508.

79. Cho HS, Mason K, Ramyar KX, Stanley AM, Gabelli SB, Denney DW, *et al*. Structure of the extracellular region of HER2 alone and in complex with the Herceptin Fab. Nature. 2003;421(6924):756–60.

80. Arteaga CL, Sliwkowski MX, Osborne CK, Perez EA, Puglisi F, Gianni L. Treatment of HER2 positive breast cancer: current status and future perspectives. Nat Rev Clin Oncol. 2011;9(1):16–32.

81. International Agency for Research on Cancer. World cancer day press release number 294: breast cancer overtakes lung cancer in terms of numbers of new cancer cases worldwide [Internet]. 2021 [cited January 27th 2021]. Available from: https://www.iarc.who.int/wp-content/uploads/2021/02/pr294_E.pdf.

82. DaCosta Byfield S, Buck PO, Blauer-Peterson C, Poston SA. ReCAP: treatment patterns and cost of care associated with initial therapy among patients diagnosed with operable

early-stage human epidermal growth factor receptor 2-overexpressed breast cancer in the United States: a real-world retrospective study. J Oncol Pract. 2016;12(2):159–67.

83. Peto R, Boreham J, Clarke M, Davies C, Beral V. UK and USA breast cancer deaths down 25% in year 2000 at ages 20-69 years. Lancet. 2000;355(9217):1822.

84. Cancer Research UK. Breast cancer survival [Internet]. 2020 [cited March 29th 2022]. Available from: https://www.cancerresearchuk.org/about-cancer/breast-cancer/survival.

85. Gennari A, Conte P, Rosso R, Orlandini C, Bruzzi P. Survival of metastatic breast carcinoma patients over a 20-year period: a retrospective analysis based on individual patient data from six consecutive studies. Cancer. 2005;104(8):1742–50.

86. Cancer Research UK. Breast cancer statistics [Internet]. 2022 [cited March 29th 2022]. Available from: https://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/breast-cancer.

87. van Loon AJ, Goldbohm RA, van den Brandt PA. Socioeconomic status and breast cancer incidence: a prospective cohort study. Int J Epidemiol. 1994;23(5):899–905.

88. Lundqvist A, Andersson E, Ahlberg I, Nilbert M, Gerdtham U. Socioeconomic inequalities in breast cancer incidence and mortality in Europe-a systematic review and meta-analysis. Eur J Public Health. 2016;26(5):804–13.

89. Clouston SAP, Manganello JA, Richards M. A life course approach to health literacy: the role of gender, educational attainment and lifetime cognitive capability. Age Ageing. 2017;46(3):493–9.

90. Akinyemiju T, Ogunsina K, Sakhuja S, Ogbhodo V, Braithwaite D. Life-course socioeconomic status and breast and cervical cancer screening: analysis of the WHO's Study on Global Ageing and Adult Health (SAGE). BMJ Open. 2016;6(11):e012753.

91. Brown D, Conway DI, McMahon AD, Dundas R, Leyland AH. Cancer mortality 1981-2016 and contribution of specific cancers to current socioeconomic inequalities in all cancer mortality: a population-based study. Cancer Epidemiol. 2021;74:102010.

92. Dreyer MS, Nattinger AB, McGinley EL, Pezzin LE. Socioeconomic status and breast cancer treatment. Breast Cancer Res Treat. 2018;167(1):1–8.

93. Lagerlund M, Bellocco R, Karlsson P, Tejler G, Lambe M. Socio-economic factors and breast cancer survival - a population-based cohort study (Sweden). Cancer Causes Control. 2005;16(4):419–30.

94. Hsieh CC, Trichopoulos D, Katsouyanni K, Yuasa S. Age at menarche, age at menopause, height and obesity as risk factors for breast cancer: associations and interactions in an international case-control study. Int J Cancer. 1990;46(5):796–800.

95. Loibl S, Poortmans P, Morrow M, Denkert C, Curigliano G. Breast cancer. Lancet. 2021;397(10286):1750–69.

96. Birring SS, Peake MD. Symptoms and the early diagnosis of lung cancer. Thorax. 2005;60(4):268–9.

97. National Institute for Health and Care Excellence. Adjuvant therapy for early and locally advanced breast cancer - NICE pathways [Internet]. 2018 [cited September 9th 2021]. Available from: https://pathways.nice.org.uk/pathways/early-and-locally-advanced-breast-cancer#path=view%3A/pathways/early-and-locally-advanced-breast-cancer/adjuvant-therapy-for-early-and-locally-advanced-breast-cancer.xml&content=view-node%3Anodes-her2-targeted-treatments.

98. Duma N, Santana-Davila R, Molina JR. Non–small cell lung cancer: epidemiology, screening, diagnosis, and treatment. Mayo Clin Proc. 2019;94(8):1623–40.

99. Cancer Research UK. Lung cancer statistics [Internet]. 2022 [cited March 29th 2022]. Available from: https://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/lung-cancer#heading-Three.

100. Ke X, Shen L. Molecular targeted therapy of cancer: the progress and future prospect. Front Lab Med. 2017;1(2):69–75.

101. Forsythe ML, Alwithenani A, Bethune D, Castonguay M, Drucker A, Flowerdew G, *et al.* Molecular profiling of non-small cell lung cancer. PLoS One. 2020;15(8):e0236580.

102. Hrubá F, Fabiáová E, Bencko V, Cassidy A, Lissowska J, Mates D, *et al.* Socioeconomic indicators and risk of lung cancer in central and eastern Europe. Cent Eur J Public Health. 2009;17(3):115–21.

103. Goussault H, Gendarme S, Assié JB, Bylicki O, Chouaïd C. Factors associated with early lung cancer mortality: a systematic review. Expert Rev Anticancer Ther. 2021;21(10):1125–33.

104. Forrest LF, Adams J, Rubin G, White M. The role of receipt and timeliness of treatment in socioeconomic inequalities in lung cancer survival: population-based, data-linkage study. Thorax. 2015;70(2):138–45.

105. Parkin DM, Boyd L, Walker LC. 16. The fraction of cancer attributable to lifestyle and environmental factors in the UK in 2010. Br J Cancer. 2011;105(Suppl 2):S77–81.

106. de Groot PM, Wu CC, Carter BW, Munden RF. The epidemiology of lung cancer. Transl Lung Cancer Res. 2018;7(3):220–33.

107. Hackshaw AK, Law MR, Wald NJ. The accumulated evidence on lung cancer and environmental tobacco smoke. BMJ. 1997;315(7114):980–8.

108. Alberg AJ, Samet JM. Epidemiology of lung cancer. Chest. 2003;123(1 Suppl):21S-49S.

109. Sun S, Schiller JH, Gazdar AF. Lung cancer in never smokers-a different disease. Nat Rev Cancer. 2007;7(10):778–90.

110. Fidler-Benaoudia MM, Torre LA, Bray F, Ferlay J, Jemal A. Lung cancer incidence in young women vs. young men: a systematic analysis in 40 countries. Int J Cancer. 2020;147(3):811–9.

111. Office for National Statistics. Life expectancy by census ward [Internet]. 2022 [cited March 27th 2022]. Available from: https://www.ons.gov.uk/visualisations/dvc479/map/index.html.

112. Mackenbach JP. Can we reduce health inequalities? An analysis of the English strategy (1997-2010). J Epidemiol Community Health. 2011;65(7):568–75.

113. Smith GD, Bartley M, Blane D. The Black Report on socioeconomic inequalities in health 10 years on. BMJ. 1990;301(6748):373–7.

114. Marmot MG, Shipley MJ, Rose G. Inequalities in death-specific explanations of a general pattern? Lancet. 1984;1(8384):1003–6.

115. Marmot MG, Smith GD, Stansfeld S, Patel C, North F, Head J, *et al.* Health inequalities among British civil servants: the Whitehall II study. Lancet Lond Engl. 1991;337(8754):1387–93.

116. Exworthy M, Blane D, Marmot M. Tackling health inequalities in the United Kingdom: the progress and pitfalls of policy. Health Serv Res. 2003;38(6 Pt 2):1905–22.

117. Barr B, Higgerson J, Whitehead M. Investigating the impact of the English health inequalities strategy: time trend analysis. BMJ. 2017;358:j3310.

118. Institute of Health Equity. Healthy equity in England: the Marmot Review 10 years on [Internet]. 2020 [cited October 7th 2020]. Available from: https://www.health.org.uk/publications/reports/the-marmot-review-10-years-on.

119. HM Government. Healthy lives, healthy people: our strategy for public health in England [Internet]. 2011 [cited March 27th 2022]. Available from: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/216096/dh_127424.pdf.

120. National Health Service. The NHS Long Term Plan [Internet]. 2019 [cited September 16th 2021]. Available from: https://www.longtermplan.nhs.uk/publication/nhs-long-term-plan/.

121. Cancer Research UK. UK health inequalities: 20,000 more cancer cases a year in the most deprived areas [Internet]. 2020 [cited April 1st 2022]. Available from: https://news.cancerresearchuk.org/2020/09/30/uk-health-inequalities-20000-more-cancer-cases-a-year-in-the-most-deprived-areas/.

122. Macmillan. Health inequalities: time to talk [Internet]. 2019 [cited September 22nd 2022]. Available from: https://www.macmillan.org.uk/assets/health-inequalities-paper-april-2019.pdf.

123. Vaccarella S, Lortet-Tieulent J, Saracci R, Fidler MM, Conway DI, Vilahur N, *et al.* Reducing social inequalities in cancer: setting priorities for research. CA Cancer J Clin. 2018;68(5):324–6.

124. Halpern MT. Cancer disparities research: it is time to come of age. Cancer. 2015;121(8):1158–9.

125. Albano JD, Ward E, Jemal A, Anderson R, Cokkinides VE, Murray T, *et al.* Cancer mortality in the United States by education level and race. J Natl Cancer Inst. 2007;99(18):1384–94.

126. Benitez Majano S, Di Girolamo C, Rachet B, Maringe C, Guren MG, Glimelius B, *et al.* Surgical treatment and survival from colorectal cancer in Denmark, England, Norway, and Sweden: a population-based study. Lancet Oncol. 2019;20(1):74–87.

127. Afshar N, English DR, Milne RL. Factors explaining socio-economic inequalities in cancer survival: a systematic review. Cancer Control. 2021;28:10732748211011956.

128. Griggs JJ. Role of nonclinical factors in the receipt of high-quality systemic adjuvant breast cancer treatment. J Clin Oncol. 2011;30(2):121–4.

129. Wallace TA, Martin DN, Ambs S. Interactions among genes, tumor biology and the environment in cancer health disparities: examining the evidence on a national and global scale. Carcinogenesis. 2011;32(8):1107–21.

130. Stein JN, Rivera MP, Weiner A, Duma N, Henderson L, Mody G, *et al.* Sociodemographic disparities in the management of advanced lung cancer: a narrative review. J Thorac Dis. 2021;13(6):3772–800.

131. Forrest L. Why are socioeconomic inequalities in receipt of treatment found for lung cancer? Lung Cancer Manag. 2013;2(3):e1001376.

132. Teng AM, Atkinson J, Disney G, Wilson N, Blakely T. Changing socioeconomic inequalities in cancer incidence and mortality: cohort study with 54 million person-years follow-up 1981-2011. Int J Cancer. 2017;140(6):1306–16.

133. Arık A, Dodd E, Cairns A, Streftaris G. Socioeconomic disparities in cancer incidence and mortality in England and the impact of age-at-diagnosis on cancer mortality. PloS One. 2021;16(7):e0253854.

134. Gadeyne S, Menvielle G, Kulhanova I, Bopp M, Deboosere P, Eikemo TA, *et al.* The turn of the gradient? Educational differences in breast cancer mortality in 18 European populations during the 2000s. Int J Cancer. 2017;141(1):33–44.

135. Palència L, Ferrando J, Marí-Dell'Olmo M, Gotsens M, Morrison J, Dzurova D, *et al.* Socio-economic inequalities on cancer mortality in nine European areas: the effect of the last economic recession. Cancer Epidemiol. 2020;69:101827.

136. Akinyemiju TF, Genkinger JM, Farhat M, Wilson A, Gary-Webb TL, Tehranifar P. Residential environment and breast cancer incidence and mortality: a systematic review and meta-analysis. BMC Cancer. 2015;15:191.

137. Singh GK, Miller BA, Hankey BF, Edwards BK. Area socioeconomic variations in US cancer incidence, mortality, stage, treatment, and survival 1975 - 1999 [Internet]. 2003 [cited April 1st 2022]. Available from: https://seer.cancer.gov/archive/publications/ses/ses_monograph.pdf.

138. National Cancer Intelligence Network and Cancer Research UK. Cancer by deprivation in England: Incidence, 1996-2010, Mortality, 1997-2011. [Internet] 2014 [cited 22nd March 2022]. Available from: http://www.ncin.org.uk/about_ncin/cancer_by_deprivation_in_england.

139. van der Heyden JHA, Schaap MM, Kunst AE, Esnaola S, Borrell C, Cox B, *et al.* Socioeconomic inequalities in lung cancer mortality in 16 European populations. Lung Cancer Amst Neth. 2009;63(3):322–30.

140. Lampert T, Kroll LE, Kuntz B, Hoebel J. Health inequalities in Germany and in international comparison: trends and developments over time. J Health Monit. 2018;20(3(supp 1)):1–24.

141. Bosetti C, Levi F, Lucchini F, Negri E, La Vecchia C. Lung cancer mortality in European women: recent trends and perspectives. Ann Oncol. 2005;16(10):1597–604.

142. Rachet B, Ellis L, Maringe C, Chu T, Nur U, Quaresma M, *et al.* Socioeconomic inequalities in cancer survival in England after the NHS cancer plan. Br J Cancer. 2010;103(4):446–53.

143. Woods LM, Rachet B, Coleman MP. Origins of socio-economic inequalities in cancer survival: a review. Ann Oncol. 2006;17(1):5–19.

144. Quaresma M, Coleman MP, Rachet B. 40-year trends in an index of survival for all cancers combined and survival adjusted for age and sex for each cancer in England and Wales, 1971-2011: a population-based study. Lancet. 2015;385(9974):1206–18.

145. Rachet B, Woods LM, Mitry E, Riga M, Cooper N, Quinn MJ, *et al.* Cancer survival in England and Wales at the end of the 20th century. Br J Cancer. 2008;99(Suppl 1):S2–10.

146. Ellis L, Coleman MP, Rachet B. How many deaths would be avoidable if socioeconomic inequalities in cancer survival in England were eliminated? A national population-based study, 1996-2006. Eur J Cancer. 2012;48(2):270–8.

147. Exarchakou A, Rachet B, Belot A, Maringe C, Coleman MP. Impact of national cancer policies on cancer survival trends and socioeconomic inequalities in England, 1996-2013: population based study. BMJ. 2018;360:k764.

148. Lyratzopoulos G, Barbiere JM, Rachet B, Baum M, Thompson MR, Coleman MP. Changes over time in socioeconomic inequalities in breast and rectal cancer survival in England and Wales during a 32-year period (1973-2004): the potential role of health care. Ann Oncol. 2011;22(7):1661–6.

149. Walters S, Maringe C, Butler J, Rachet B, Barrett-Lee P, Bergh J, *et al.* Breast cancer survival and stage at diagnosis in Australia, Canada, Denmark, Norway, Sweden and the UK, 2000-2007: a population-based study. Br J Cancer. 2013;108(5):1195–208.

150. Ingleby FC, Woods LM, Atherton IM, Baker M, Elliss-Brookes L, Belot A. An investigation of cancer survival inequalities associated with individual-level socio-economic status, area-level deprivation, and contextual effects, in a cancer patient cohort in England and Wales. BMC Public Health. 2022;22(1):90.

151. Redondo-Sánchez D, Petrova D, Rodríguez-Barranco M, Fernández-Navarro P, Jiménez-Moleón JJ, Sánchez MJ. Socio-economic inequalities in lung cancer outcomes: an overview of systematic reviews. Cancers. 2022;14(2):398.

152. Pollock AM, Vickers N. Breast, lung and colorectal cancer incidence and survival in South Thames region, 1987-1992: the effect of social deprivation. J Public Health Med. 1997;19(3):288–94.

153. Rich AL, Tata LJ, Stanley RA, Free CM, Peake MD, Baldwin DR, *et al.* Lung cancer in England: information from the National Lung Cancer Audit (LUCADA). Lung Cancer. 2011;72(1):16–22.

154. Finke I, Behrens G, Weisser L, Brenner H, Jansen L. Socioeconomic differences and lung cancer survival - systematic review and meta-analysis. Front Oncol. 2018;8:536.

155. Mehta S, Shelling A, Muthukaruppan A, Lasham A, Blenkiron C, Laking G, *et al.* Predictive and prognostic molecular markers for cancer medicine. Ther Adv Med Oncol. 2010;2(2):125–48.

156. Shin SH, Bode AM, Dong Z. Precision medicine: the foundation of future cancer therapeutics. NPJ Precis Oncol. 2017;1(1):12.

157. Faguet GB. A brief history of cancer: age-old milestones underlying our current knowledge database. Int J Cancer. 2015;136(9):2022–36.

158. Huang M, Shen A, Ding J, Geng M. Molecularly targeted cancer therapy: some lessons from the past decade. Trends Pharmacol Sci. 2014;35(1):41–50.

159. Summerhayes M. What targeted therapy for cancer is and how it is achieved: a review. Pharm J. 2006;277:518–25.

160. Hait WN, Hambley TW. Targeted cancer therapeutics. Cancer Res. 2009;69(4):1263–7; discussion 1267.

161. Lehmann F, Wennerberg J. Evolution of nitrogen-based alkylating anticancer agents. Processes. 2021;9(2):377.

162. Kaldor J. Second cancer following chemotherapy and radiotherapy. An epidemiological perspective. Acta Oncol. 1990;29(5):647–55.

163. Volkova M, Russell R. Anthracycline cardiotoxicity: prevalence, pathogenesis and treatment. Curr Cardiol Rev. 2011;7(4):214–20.

164. Gianfaldoni S, Gianfaldoni R, Wollina U, Lotti J, Tchernev G, Lotti T. An overview on radiotherapy: from its history to its current applications in dermatology. Open Access Maced J Med Sci. 2017;5(4):521–5.

165. Cancer Research UK. What is radiotherapy? [Internet]. 2020 [cited June 20th 2020]. Available from: https://www.cancerresearchuk.org/about-cancer/cancer-in-general/treatment/radiotherapy/about.

166. National Institute for Health and Care Excellence. Recommendations, early and locally advanced breast cancer: diagnosis and management guidance [Internet]. 2018 [cited March 30th 2022]. Available from:

https://www.nice.org.uk/guidance/ng101/chapter/Recommendations#surgery-to-the-breast.

167. Tosello G, Torloni MR, Mota BS, Neeman T, Riera R. Breast surgery for metastatic breast cancer. Cochrane Database Syst Rev. 2018;2018(3):CD011276.

168. National Institute for Health and Care Excellence. Recommendations, advanced breast cancer: diagnosis and treatment, guidance [Internet]. 2017 [cited March 30th 2022]. Available from: https://www.nice.org.uk/guidance/cg81/chapter/Recommendations/.

169. Jutzy JMS, Lemons JM, Luke JJ, Chmura SJ. The evolution of radiation therapy in metastatic breast cancer: from local therapy to systemic agent. Int J Breast Cancer. 2018;2018:4786819.

170. Zappa C, Mousa SA. Non-small cell lung cancer: current treatment and future advances. Transl Lung Cancer Res. 2016;5(3):288–300.

171. David EA, Clark JM, Cooke DT, Melnikow J, Kelly K, Canter RJ. The role of thoracic surgery in the therapeutic management of metastatic non-small cell lung cancer. J Thorac Oncol. 2017;12(11):1636–45.

172. Sabatello M. Precision medicine, health disparities, and ethics: the case for disability inclusion. Genet Med. 2018;20(4):397–9.

173. Hanahan D, Weinberg RA. Hallmarks of cancer: the next generation. Cell. 2011;144(5):646–74.

174. Seshacharyulu P, Ponnusamy MP, Haridas D, Jain M, Apar KG, Batra SK. Targeting the EGFR signaling pathway in cancer therapy. Expert Opin Ther Targets. 2012;16(1):15–31.

175. Collins FS, Varmus H. A new initiative on precision medicine. N Engl J Med. 2015;372(9):793–5.

176. Prosperi M, Min JS, Bian J, Modave F. Big data hurdles in precision medicine and precision public health. BMC Med Inform Decis Mak. 2018;18(1):139.

177. Khoury MJ, Galea S. Will precision medicine improve population health? JAMA. 2016;316(13):1357–8.

178. Jain KK. From molecular diagnostics to personalized medicine. The IBC workshop, London, UK, 1st May, 2002. Expert Rev Mol Diagn. 2002;2(4):299–301.

179. Pusztai L. Is precision medicine ready for use in breast cancer? Clin Adv Hematol Oncol. 2016;14(12):964–6.

180. Jaffe S. Planning for US precision medicine initiative underway. The Lancet. 2015;385(9986):2448–9.

181. Hayes DF, Markus HS, Leslie RD, Topol EJ. Personalized medicine: risk prediction, targeted therapies and mobile health technology. BMC Med. 2014;12(1):37.

182. Prasad V. Perspective: the precision-oncology illusion. Nature. 2016;537(7619):S63–S63.

183. Brock A, Huang S. Precision oncology: between vaguely right and precisely wrong. Cancer Res. 2017;77(23):6473–9.

184. Pritchard KI. Endocrine therapy: is the first generation of targeted drugs the last? J Intern Med. 2013;274(2):144–52.

185. Ross J. Targeted therapies for cancer. In: Gingsburg GS and Willard HF, editors. Essentials of genomic and personalized medicine. Amsterdam Boston London: Elsevier; 2010. p. 532–47.

186. National Cancer Institute. National Cancer Institute dictionaries [Internet]. 2022 [cited March 30th 2022]. Available from: https://www.cancer.gov/publications/dictionaries/cancer-terms/def/biomarker?redirect=true.

187. Butts C, Kamel-Reid S, Batist G, Chia S, Blanke C, Moore M, *et al.* Benefits, issues, and recommendations for personalized medicine in oncology in Canada. Curr Oncol. 2013;20(5):e475-483.

188. Schilsky RL. Drivers of change in cancer care. J Oncol Pract. 2014;10(5):315–6.

189. Tuckson RV, Newcomer L, De Sa JM. Accessing genomic medicine: affordability, diffusion, and disparities. JAMA. 2013;309(14):1469–70.

190. Nadal E, Massuti B, Dómine M, García-Campelo R, Cobo M, Felip E. Immunotherapy with checkpoint inhibitors in non-small cell lung cancer: insights from long-term survivors. Cancer Immunol Immunother. 2019;68(3):341–52.

191. Gray M, Lagerberg T, Dombrádi V. Equity and value in 'precision medicine'. New Bioeth. 2017;23(1):87–94.

192. Haber DA, Gray NS, Baselga J. The evolving war on cancer. Cell. 2011;145(1):19–24.

193. Adizie JB, Tweedie J, Khakwani A, Peach E, Hubbard R, Wood N, *et al.* Biomarker testing for people with advanced lung cancer in England. JTO Clin Res Rep. 2021;2(6):100176.

194. National Institute for Health and Care Excellence. Lung cancer: diagnosis and management NICE Guidance [Internet]. 2022 [cited February 10th 2022]. Available from: https://www.nice.org.uk/guidance/ng122/chapter/Context.

195. National Health Service England. National genomic test directory for cancer 2021/2022 [Internet]. 2021 [cited April 2nd 2022]. Available from: https://www.england.nhs.uk/publication/national-genomic-test-directories/.

196. Evans M, O'Sullivan B, Smith M, Hughes F, Mullis T, Trim N, *et al.* Large-scale EGFR mutation testing in clinical practice: analysis of a series of 18,920 non-small cell lung cancer cases. Pathol Oncol Res. 2019;25(4):1401–9.

197. Arbour KC, Riely GJ. Systemic therapy for locally advanced and metastatic non-small cell lung cancer: a review. JAMA. 2019;322(8):764–74.

198. National Cancer Institute. Molecularly targeted therapy definition [Internet]. 2022 [cited April 4th 2022]. Available from: https://www.cancer.gov/publications/dictionaries/cancer-terms/def/molecularly-targeted-therapy.

199. Greystoke A, Chaturvedi A. An introduction to stratified medicine. Drug Discov Today. 2015 Dec;20(12):1409–13.

200. Cancer Research UK. What is immunotherapy [Internet]. 2021 [cited April 4th 2022]. Available from: https://www.cancerresearchuk.org/about-cancer/cancer-in-general/treatment/immunotherapy/what-is-immunotherapy.

201. Davar D, Kirkwood JM. PD-1 immune checkpoint inhibitors and immune related adverse events: understanding the upside of the downside of checkpoint blockade. JAMA Oncol. 2019;5(7):942–3.

202. European Society for Medical Oncology. ESMO clinical practice guidelines: lung and chest tumours [Internet]. 2022 [cited April 3rd 2022]. Available from: https://www.esmo.org/guidelines/lung-and-chest-tumours.

203. American Society of Clinical Oncology. ASCO guidelines, tools & resources [Internet]. 2022 [cited March 22nd 2022]. Available from: https://www.asco.org/practice-patients/guidelines.

204. Cancer Research UK. Immunotherapy for triple negative breast cancer approved for NHS in England [Internet]. 2020 [cited April 4th 2022]. Available from: https://news.cancerresearchuk.org/2020/05/26/immunotherapy-for-triple-negative-breast-cancer-approved-for-nhs-in-england/.

205. Burstein HJ. Patient education: treatment of early HER2-positive breast cancer (beyond the basics) [Internet]. 2022 [cited April 2nd 2022]. Available from: https://www.uptodate.com/contents/treatment-of-early-her2-positive-breast-cancer-beyond-the-basics/print.

206. Capelan M, Pugliano L, De Azambuja E, Bozovic I, Saini KS, Sotiriou C, *et al.* Pertuzumab: new hope for patients with HER2 positive breast cancer. Ann Oncol. 2013;24(2):273–82.

207. National Institute for Health and Care Excellence. Osimertinib for adjuvant treatment of EGFR mutation-positive non-small-cell lung cancer after complete tumour resection NICE guidance [Internet]. 2022 [cited December 3rd 2022]. Available from: https://www.nice.org.uk/guidance/ta761.

208. Remon J, Soria JC, Peters S. eUpdate – early and locally advanced non-small-cell lung cancer [Internet]. 2021 [cited June 21st 2022]. Available from: https://www.esmo.org/guidelines/guidelines-by-topic/lung-and-chest-tumours/early-stage-and-locally-advanced-non-metastatic-non-small-cell-lung-cancer/eupdate-early-and-locally-advanced-non-small-cell-lung-cancer-nsclc-treatment-recommendations2.

209. Planchard D, Popat S, Kerr K, Novello S, Smit EF, Faivre-Finn C, *et al.* Metastatic non-small-cell lung cancer: ESMO clinical practice guidelines for diagnosis, treatment and follow-up. Ann Oncol 2018;29(Suppl 4):iv192–237.

210. Planchard D, Popat S, Kerr K, Novello S, Smit EF, Faivre-Finn, *et al.* Clinical practice living guidelines - metastatic non-small-cell lung cancer, ESMO (updated version) [Internet]. 2020 [cited June 21st 2022]. Available from: https://www.esmo.org/guidelines/guidelines-by-topic/lung-and-chest-tumours/clinical-practice-living-guidelines-metastatic-non-small-cell-lung-cancer.

211. National Institute for Health and Care Excellence. Lung cancer update: development of a systemic anti-cancer therapy algorithm. NICE guideline NG122 [Internet]. 2019 [cited April 20th 2022]. Available from: https://www.nice.org.uk/guidance/ng122/evidence/e-development-of-a-systemic-anticancer-therapy-algorithm-pdf-6722112210.

212. Baxevanos P, Mountzios G. Novel chemotherapy regimens for advanced lung cancer: have we reached a plateau? Ann Transl Med. 2018;6(8):139.

213. Li R, Daniel R, Rachet B. How much do tumor stage and treatment explain socioeconomic inequalities in breast cancer survival? Applying causal mediation analysis to population-based data. Eur J Epidemiol. 2016;31:603–11.

214. European Cancer Organisation. It can be done - beating inequalities in cancer care action report [Internet]. 2020 [cited February 28th 2022]. Available from: https://www.europeancancer.org/resources/164:beating-inequalities-in-cancer-care.html.

215. All Party Parliamentary Group on Cancer. Report of the All Party Parliamentary Group on Cancer's inquiry into inequalities in cancer [Internet]. 2009 [cited February 28th 2022]. Available from: https://www.macmillan.org.uk/documents/getinvolved/campaigns/appg/britainagainstcancer2009/cancerinequalitiesreport.pdf.

216. Byers TE, Wolf HJ, Bauer KR, Bolick-Aldrich S, Chen VW, Finch JL, *et al.* The impact of socioeconomic status on survival after cancer in the United States: findings from the national program of cancer registries patterns of care study. Cancer. 2008;113(3):582–91.

217. Michalski TA, Nattinger AB. The influence of black race and socioeconomic status on the use of breast-conserving surgery for Medicare beneficiaries. Cancer. 1997;79(2):314–9.

218. Smith GL, Xu Y, Shih YCT, Giordano SH, Smith BD, Hunt KK, *et al.* Breast-conserving surgery in older patients with invasive breast cancer: current patterns of treatment across the United States. J Am Coll Surg. 2009;209(4):425-433.e2.

219. Raine R, Wong W, Scholes S, Ashton C, Obichere A, Ambler G. Social variations in access to hospital care for patients with colorectal, breast, and lung cancer between 1999 and 2006: retrospective analysis of hospital episode statistics. BMJ. 2010;340:b5479.

220. Neuner JM, Kong A, Blaes A, Riley D, Chrischilles E, Smallwood A, *et al.* The association of socioeconomic status with receipt of neoadjuvant chemotherapy. Breast Cancer Res Treat. 2019;173(1):179–88.

221. Forrest LF, Adams J, Wareham H, Rubin G, White M. Socioeconomic inequalities in lung cancer treatment: systematic review and meta-analysis. PLoS Med. 2013;10(2):e1001376.

222. Ebner PJ, Ding L, Kim AW, Atay SM, Yao MJ, Toubat O, *et al.* The effect of socioeconomic status on treatment and mortality in non-small cell lung cancer patients. Ann Thorac Surg. 2020;109(1):225–32.

223. Forrest LF, White M, Rubin G, Adams J. The role of patient, tumour and system factors in socioeconomic inequalities in lung cancer treatment: population-based study. Br J Cancer. 2014;111(3):608–18.

224. Jones AP, Haynes R, Sauerzapf V, Crawford SM, Zhao H, Forman D. Travel time to hospital and treatment for breast, colon, rectum, lung, ovary and prostate cancer. Eur J Cancer. 2008;44(7):992–9.

225. Bayer R, Galea S. Public health in the precision-medicine era. N Engl J Med. 2015;373(6):499–501.

226. Bilkey GA, Burns BL, Coles EP, Mahede T, Baynam G, Nowak KJ. Optimizing precision medicine for public health. Front Public Health. 2019;7:7-42.

227. Armstrong K. Genomics and health care disparities: the role of statistical discrimination. JAMA. 2012;308(19):1979–80.

228. Geneviève LD, Martani A, Shaw D, Elger BS, Wangmo T. Structural racism in precision medicine: leaving no one behind. BMC Med Ethics. 2020;21(1):17.

229. Matthew DB. Two threats to precision medicine equity. Ethn Dis. 2019;29(Suppl 3):629–40.

230. Victora CG, Vaughan JP, Barros FC, Silva AC, Tomasi E. Explaining trends in inequities: evidence from Brazilian child health studies. Lancet. 2000;356(9235):1093–8.

231. Norris RP, Dew R, Sharp L, Greystoke A, Rice S, Johnell K, *et al.* Are there socio-economic inequalities in utilization of predictive biomarker tests and biological and precision therapies for cancer? A systematic review and meta-analysis. BMC Med. 2020; 18(1):282.

232. Dranitsaris G, Zhu X, Adunlin G, Vincent MD. Cost effectiveness vs. affordability in the age of immuno-oncology cancer drugs. Expert Rev Pharmacoecon Outcomes Res. 2018;18(4):351–7.

233. National Comprehensive Cancer Network. Clinical practice guidelines in oncology: non-small cell lung cancer [Internet]. 2020 [cited September 22nd 2020]. Available from: https://www.nccn.org/store/login/login.aspx?ReturnURL=https://www.nccn.org/professionals/physician_gls/pdf/nscl.pdf.

234. IQVIA Institute. Global oncology trends 2019 therapeutics, clinical development and health system implications institute report [Internet]. 2019 [cited September 22nd 2020]. Available from: https://www.iqvia.com/insights/the-iqvia-institute/reports/global-oncology-trends-2019.

235. IQVIA Institute. Global oncology trends 2021 - outlook to 2025 [Internet]. 2021 [cited January 26th 2026]. Available from: https://www.iqvia.com/insights/the-iqvia-institute/reports/global-oncology-trends-2021.

236. Martin AP, Downing J, Cochrane M, Collins B, Francis B, Haycox A, *et al.* Trastuzumab uptake in HER2-positive breast cancer patients: a systematic review and meta-analysis of observational studies. Crit Rev Oncol Hematol. 2018;130:92–107.

237. McClellan KA, Avard D, Simard J, Knoppers BM. Personalized medicine and access to health care: potential for inequitable access? Eur J Hum Genet. 2013;21(2):143–7.

238. Ward MM. Personalized therapeutics: a potential threat to health equity. J Gen Intern Med. 2012;27(7):868–70.

239. Huey RW, Hawk E, Offodile AC. Mind the gap: precision oncology and its potential to widen disparities. J Oncol Pract. 2019;15(6):301–4.

240. Schwaederle M, Zhao M, Lee JJ, Eggermont AM, Schilsky RL, Mendelsohn J, *et al.* Impact of precision medicine in diverse cancers: a meta-analysis of phase II clinical trials. J Clin Oncol. 2015;33(32):3817–25.

241. Mok TS, Wu YL, Thongprasert S, Yang CH, Chu DT, Saijo N, *et al.* Gefitinib or carboplatin-paclitaxel in pulmonary adenocarcinoma. N Engl J Med. 2009;361(10):947–57.

242. Moher D, Liberati A, Tetzlaff J, Altman DG, Group TP. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. PLoS Med. 2009;6(7):e1000097.

243. Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, *et al.* The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ. 2021;372:n71.

244. Vickers E. A beginner's guide to targeted cancer treatments. Oxford: Wiley Blackwell; 2018.

245. European Medicines Agency. European public assessment reports for human and veterinary medicines spreadsheet [Internet]. 2019 [cited January 2nd 2019]. Available from: https://www.ema.europa.eu/en/medicines/download-medicine-data.

246. Joint Formulary Committee. BNF 76: September 2018 – March 2019. London: Pharmaceutical Press; 2018.

247. Higgins J, Deek J. Chapter 7: selecting studies and collecting data. In: Higgins JPT and Green S, editors. Cochrane handbook for systematic reviews of interventions version 510 [updated March 2011]. Chichester: Wiley; 2011.

248. Motheral B, Brooks J, Clark MA, Crown WH, Davey P, Hutchins D, *et al*. A checklist for retrospective database studies - report of the ISPOR task force on retrospective databases. Value Health. 2003;6(2):90-97.

249. Jaksa A, Wu J, Jónsson P, Eichler HG, Vititoe S, Gatto NM. Organized structure of real-world evidence best practices: moving from fragmented recommendations to comprehensive guidance. J Comp Eff Res. 2021;10(9):711–31.

250. Wells G, Shea B, O'Connell, D, Peterson J, Welch V, Losos M, *et al.* The Newcastle-Ottawa Scale (NOS) for assessing the quality of nonrandomised studies in meta-analyses. [Internet]. 2013 [cited December 5th 2022]. Available from: https://www.ohri.ca/programs/clinical_epidemiology/oxford.asp.

251. Toye F, Seers K, Allcock N, Briggs M, Carr E, Andrews J, *et al.* 'Trying to pin down jelly' - exploring intuitive processes in quality assessment for meta-ethnography. BMC Med Res Methodol. 2013;13:46.

252. Murphy E, Dingwall R, Greatbatch D, Parker S, Watson P. Qualitative research methods in health technology assessment: a review of the literature. Health Technol Assess. 1998;2(16):iii–ix, 1–274.

253. Dixon-Woods M, Sutton A, Shaw R, Miller T, Smith J, Young B, *et al*. Appraising qualitative research for inclusion in systematic reviews: a quantitative and qualitative comparison of three methods. J Health Serv Res Policy. 2007;12(1):42–7.

254. Deeks, JJ, Higgins, JPT, Altman DG. Chapter 9: analysing data and undertaking meta-analyses: sensitivity analyses, section 9.7. In: Higgins JPT and Green S, editors. Cochrane handbook for systematic reviews of interventions (version 510 [updated March 2011]). Chichester: John Wiley & Sons; 2011.

255. Deeks J, Higgins J, Altman D. Chapter 10: analysing data and undertaking meta-analyses, section 4.6.4. In: Higgins JPT and Green S, editors. Cochrane handbook for systematic reviews of interventions (version 510 [updated March 2011]). Chichester: John Wiley & Sons; 2011.

256. National Cancer Institute. SEER registries - about the SEER registry [Internet]. SEER. 2022 [cited June 22nd 2022]. Available from: https://seer.cancer.gov/registries/index.html.

257. Deeks JJ, Higgins JPT, Altman DG. Chapter 9: analysing data and undertaking meta-analyses: sensitivity analyses, section 9.7. In: Higgins JPT and Green S, editors. Cochrane handbook for systematic reviews of interventions (version 510 [updated March 2011]). Chichester: John Wiley & Sons; 2011.

258. Sharpe D. Of apples and oranges, file drawers and garbage: why validity issues in meta-analysis will not go away. Clin Psychol Rev. 1997;17(8):881–901.

259. Guzzo RA, Jackson SF, Katzell RA. Meta-analysis analysis. Res Organ Behav. 1987;9:407–42.

260. Feeley TH. Assessing study quality in meta-analysis. Hum Commun Res. 2020;46(2–3):334–42.

261. Eysenck HJ. Meta-analysis: an abuse of research integration. J Spec Educ. 1984;18(1):41–59.

262. Rosenthal R. An evaluation of procedures and results. In: Wachter KW and Straf ML, editors. The future of meta-analysis. Russel Sage Foundation; 1990. p. 123–33.

263. Eysenck HJ. An exercise in mega-silliness. Am Psychol. 1978;33(5):517.

264. Higgins JPT, Thompson SG. Quantifying heterogeneity in a meta-analysis. Stat Med. 2002;21(11):1539–58.

265. Deeks J, Higgins J, Altman D. Identifying and measuring heterogeneity. In: Higgins JPT and Green S, editors. Cochrane handbook for systematic reviews of interventions, (version 510 [updated March 2011]). Chichester: John Wiley & Sons; 2011.

266. Guyatt GH, Oxman AD, Kunz R, Woodcock J, Brozek J, Helfand M, *et al.* GRADE guidelines 7 rating the quality of evidence-inconsistency. J Clin Epidemiol. 2011;64(12):1294–302.

267. Egger M, Davey Smith G, Schneider M, Minder C. Bias in meta-analysis detected by a simple, graphical test. BMJ. 1997;315(7109):629–34.

268. Freedman RA, Vaz-Luis I, Barry WT, Lii H, Lin NU, Winer EP, *et al.* Patterns of chemotherapy, toxicity, and short-term outcomes for older women receiving adjuvant trastuzumab-based therapy. Breast Cancer Res Treat. 2014;145(2):491–501.

269. Vaz-Luis I, Keating NL, Lin NU, Lii H, Winer EP, Freedman RA. Duration and toxicity of adjuvant trastuzumab in older patients with early-stage breast cancer: a population-based study. J Clin Oncol. 2014;32(9):927–34.

270. Reeder-Hayes KE, Meyer AM, Hinton SP, Meng K, Carey LA, Dusetzina SB. Comparative toxicity and effectiveness of trastuzumab-based chemotherapy regimens in older women with early-stage breast cancer. J Clin Oncol. 2017;35(29):3298–305.

271. Chavez-MacGregor M, Niu J, Zhang N, Elting LS, Smith BD, Banchs J, *et al.* Cardiac monitoring during adjuvant trastuzumab-based chemotherapy among older patients with breast cancer. J Clin Oncol. 2015;33(19):2176–83.

272. Lu CY, Srasuebkul P, Drew AK, Chen K, Ward RL, Pearson SA. Trastuzumab therapy in Australia: which patients with HER2+ metastatic breast cancer are assessed for cardiac function? Breast. 2013;22(4):482–7.

273. Krimphove MJ, Tully KH, Friedlander DF, Marchese M, Ravi P, Lipsitz SR, *et al.* Adoption of immunotherapy in the community for patients diagnosed with metastatic melanoma. J Immunother Cancer. 2019;7(1):289.

274. Zheng Z, Hanna N, Onukwugha E, Reese ES, Seal B, Mullins CD. Does the type of first-line regimens influence the receipt of second-line chemotherapy treatment? An analysis of 3211 metastatic colon cancer patients. Cancer Med. 2014;3(1):124–33.

275. Kwan SW, Harris WP, Gold LS, Hebert PL. Comparative effectiveness of transarterial embolization and sorafenib for hepatocellular carcinoma: a population-based study. Am J Roentgenol. 2018;210(6):1359–65.

276. Li P, Jahnke J, Pettit AR, Wong YN, Doshi JA. Comparative survival associated with use of targeted vs nontargeted therapy in Medicare patients with metastatic renal cell carcinoma. JAMA Netw Open. 2019;2(6).

277. Zandberg DP, Cullen K, Bentzen SM, Goloubeva OG. Definitive radiation with concurrent cetuximab vs. radiation with or without concurrent cytotoxic chemotherapy in older patients with squamous cell carcinoma of the head and neck: Analysis of the SEER-Medicare linked database. Oral Oncol. 2018;86:132–40.

278. Pensa M, Swede H, Brockmeyer JA, Gregorio DI. Patterns of HER2 testing in the management of primary breast cancer. Cancer Epidemiol. 2009;33(2):113–7.

279. Lund MJ, Butler EN, Hair BY, Ward KC, Andrews JH, Oprea-Ilies G, *et al.* Age/race differences in HER2 testing and in incidence rates for breast cancer triple subtypes: a population-based study and first report. Cancer. 2010;116(11):2549–59.

280. Ferrusi IL, Earle CC, Trudeau M, Leighl NB, Pullenayegum E, Khong H, *et al.* Closing the personalized medicine information gap: HER2 test documentation practice. Am J Manag Care. 2013;19(1):838–44.

281. de Camargo Cancela M, Comber H, Sharp L. Which women with breast cancer do, and do not, undergo receptor status testing? A population-based study. Cancer Epidemiol. 2015;39(5):778–82.

282. Greenbaum A, Wiggins C, Meisner AL, Rojo M, Kinney AY, Rajput A. KRAS biomarker testing disparities in colorectal cancer patients in New Mexico. Heliyon. 2017;3(11):e00448.

283. Rico A, Pollack LA, Thompson TD, Hsieh M-C, Wu X-C, Karlitz JJ, *et al.* KRAS testing and first-line treatment among patients diagnosed with metastatic colorectal cancer using population data from ten national program of cancer registries in the United States. J Cancer Res Ther (Manch). 2016;5(2):7–13.

284. Webster J, Kauffman TL, Feigelson HS, Pawloski PA, Onitilo AA, Potosky AL, *et al.* KRAS testing and epidermal growth factor receptor inhibitor treatment for colorectal cancer in community settings. Cancer Epidemiol Biomarkers Prev. 2013;22(1):91–101.

285. Presley CJ, Tang D, Soulos PR, Chiang AC, Longtine JA, Adelson KB, *et al.* Association of broad-based genomic sequencing with survival among patients with advanced non-small cell lung cancer in the community oncology setting. JAMA. 2018;320(5):469–77.

286. Du XL, Xia R, Burau K, Liu CC. Cardiac risk associated with the receipt of anthracycline and trastuzumab in a large nationwide cohort of older women with breast cancer, 1998–2005. Med Oncol. 2011;28(1):80–90.

287. Vaz-Luis I, Lin NU, Keating NL, Barry WT, Lii H, Winer EP, *et al.* Racial differences in outcomes for patients with metastatic breast cancer by disease subtype. Breast Cancer Res Treat. 2015;151(3):697–707.

288. Reeder-Hayes K, Peacock Hinton S, Meng K, Carey LA, Dusetzina SB. Disparities in use of human epidermal growth hormone receptor 2–targeted therapy for early-stage breast cancer. J Clin Oncol. 2016;34(17):2003–9.

289. Vaz-Luis I, Lin NU, Keating NL, Barry WT, Lii J, Burstein HJ, *et al.* Treatment of early-stage human epidermal growth factor 2-positive cancers among medicare enrollees: age and race strongly associated with non-use of trastuzumab. Breast Cancer Res Treat. 2016;159(1):151–62.

290. Tsai HT, Isaacs C, Lynce FC, O'Neill SC, Liu C, Schwartz MD, *et al.* Initiation of trastuzumab by women younger than 64 years for adjuvant treatment of stage I–III breast cancer. J Natl Compr Cancer Netw. 2017;15(5):601–7.

291. Freedman RA, Hughes ME, Ottesen RA, Weeks JC, He Y, Wong YN, *et al.* Use of adjuvant trastuzumab in women with human epidermal growth factor receptor 2 (HER2)-positive breast cancer by race/ethnicity and education within the National Comprehensive Cancer Network. Cancer. 2013;119(4):839–46.

292. Haas JS, Phillips KA, Liang SY, Hassett MJ, Keohane C, Elkin EB, *et al.* Genomic testing and therapies for breast cancer in clinical practice. J Oncol Pract. 2011;1(17):e174-81.

293. Goldhar HA, Yan AT, Ko DT, Earle CC, Tomlinson GA, Trudeau ME, *et al.* The temporal risk of heart failure associated with adjuvant trastuzumab in breast cancer patients: a population study. J Natl Cancer Inst. 2016;108(1).

294. Kumachev A, Trudeau ME, Chan KKW. Associations among socioeconomic status, patterns of care and outcomes in breast cancer patients in a universal health care system: Ontario's experience. Cancer. 2016;122(6):893–8.

295. Thavendiranathan P, Abdel-Qadir H, Fischer HD, Camacho X, Amir E, Austin PC, *et al.* Breast cancer therapy-related cardiac dysfunction in adult women treated in routine clinical practice: a population-based cohort study. J Clin Oncol. 2016;34(19):2239–46.

296. Li J, Shao Z, Xu B, Jiang Z, Cui S, Zhang J, *et al.* Use of trastuzumab as an adjuvant/neoadjuvant therapy in patients with HER2-positive breast cancer in China: the Nvwa study. Medicine (Baltimore). 2018;97(21):e10350.

297. Shih Y-CT, Elting LS, Halpern MT. Factors associated with immunotherapy use among newly diagnosed cancer patients. Med Care. 2009;47(9):948–58.

298. Zhu J, Sharma DB, Gray SW, Chen AB, Weeks JC, Schrag D. Carboplatin and paclitaxel with vs without bevacizumab in older patients with advanced non-small cell lung cancer. JAMA. 2012;307(15):1593–601.

299. Langer C, Ravelo A, Hazard SJ, Guerin A, Ionescu-Ittu R, Latremouille-Viau D, *et al.* Comparison of survival and hospitalization rates between Medicare patients with advanced NSCLC treated with bevacizumab-carboplatin-paclitaxel and carboplatin-paclitaxel: a retrospective cohort study. Lung Cancer. 2014;86(3):350–7.

300. Ritzwoller DP, Carroll NM, Delate T, Hornbrook MC, Kushi L, Bowles EJA, *et al.* Comparative effectiveness of adjunctive bevacizumab for advanced lung cancer: the cancer research network experience. J Thorac Oncol. 2014;9(5):692–701.

301. Menter AR, Carroll NM, Sakoda LC, Delate T, Hornbrook MC, Jain RK, *et al.* Effect of angiotensin system inhibitors on survival in patients receiving chemotherapy for advanced non-small-cell lung cancer. Clin Lung Cancer. 2017;18(2):189-197.e3.

302. Maguire FB, Morris CR, Parikh-Patel A, Cress RD, Keegan THM, Li CS, *et al*. First-line systemic treatments for stage IV non-small cell lung cancer in California: patterns of care and outcomes in a real-world setting. JNCI Cancer Spectr. 2019b;3(3):pkz020.

303. Verma V, Haque W, Cushman TR, Lin C, Simone CB, Chang JY, *et al.* Racial and insurance-related disparities in delivery of immunotherapy-type compounds in the United States. J Immunother. 2019;42(2):55–64.

304. Lairson D, Parikh R, Cormier J, Chan W, Du X. Cost-utility analysis of treatments for advanced non-small cell lung cancer. Am J Pharm Benefits. 2015;7(6):271–9.

305. Sahara K, Farooq SA, Tsilimigras DI, Merath K, Paredes AZ, Wu L, *et al.* Immunotherapy utilization for hepatobiliary cancer in the United States: disparities among patients with different socioeconomic status. Hepatobiliary Surg Nutr. 2019;9(1):13–24.

306. Sanoff HK, Chang Y, Lund JL, O'Neil BH, Dusetzina SB. Sorafenib effectiveness in advanced hepatocellular carcinoma. The Oncologist. 2016;21(9):1113–20.

307. Parsons HM, Chu Q, Karlitz JJ, Stevens JL, Harlan LC. Adoption of sorafenib for the treatment of advanced-stage hepatocellular carcinoma in oncology practices in the United States. Liver Cancer. 2017;6(3):216–26.

308. Sarpel U, Heskel M, Spivack JH, Feferman Y, Ang C, Gany F. Disparities in access to sorafenib in communities with low socioeconomic status. J Health Care Poor Underserved. 2018;29(3):1123–34.

309. Al-Qurayshi Z, Crowther JE, Hamner JB, Ducoin C, Killackey MT, Kandil E. Disparities of immunotherapy utilization in patients with stage III cutaneous melanoma: a national perspective. Anticancer Res. 2018;38(5):2897–901.

310. Haque W, Verma V, Butler EB, Teh BS. Racial and socioeconomic disparities in the delivery of immunotherapy for metastatic melanoma in the United States. J Immunother. 2019;42(6):228–35.

311. Fu AZ, Tsai HT, Marshall JL, Freedman AN, Potosky AL. Utilization of bevacizumab in US elderly patients with colorectal cancer receiving chemotherapy. J Oncol Pharm Pract. 2014;20(5):332–40.

312. Meyerhardt JA, Li L, Sanoff HK, Carpenter W, Schrag D. Effectiveness of bevacizumab with first-line combination chemotherapy for Medicare patients with stage IV colorectal cancer. J Clin Oncol. 2012;30(6):608–15.

313. Cen P, Liu C, Du XL. Comparison of toxicity profiles of fluorouracil versus oxaliplatin regimens in a large population-based cohort of elderly patients with colorectal cancer. Ann Oncol. 2012;23(6):1503–11.

314. Parikh RC, Du XL, Morgan RO, Lairson DR. Patterns of treatment sequences in chemotherapy and targeted biologics for metastatic colorectal cancer: findings from a large community-based cohort of elderly patients. Drugs Real World Outcomes. 2016;3(1):69–82.

315. Neugut AI, Becker DJ, Insel BJ, Hershman DL. Uptake of oxaliplatin and bevacizumab for treatment of node-positive and metastatic colon cancer. J Oncol Pract. 2012;8(3):156–63.

316. Raab GT, Lin A, Hillyer GC, Keller D, O'Neil DS, Accordino MK, *et al.* Use of bevacizumab for elderly patients with stage IV colon cancer: analysis of SEER-Medicare data. Clin Colorectal Cancer. 2019;18(3):e294–9.

317. Taylor JP, Stem M, Yu D, Chen SY, Fang SH, Gearhart SL, *et al.* Treatment strategies and survival trends for anorectal melanoma: is it time for a change? World J Surg. 2019;43(7):1809–19.

318. Saigal C, Deibert C, Lai J, Schonlau M. Disparities in the treatment of patients with IL-2 for metastatic renal cell carcinoma. Urol Oncol. 2010;28(3):308–13.

319. Amini A, Eguchi M, Jones BL, Stokes WA, Gupta A, McDermott JD, *et al.* Comparing outcomes of concurrent chemotherapy regimens in patients 65 years or older with locally advanced oropharyngeal carcinoma. Cancer. 2018;124(22):4322–31.

320. Xiang M, Holsinger FC, Colevas AD, Chen MM, Le QT, Beadle BM. Survival of patients with head and neck cancer treated with definitive radiotherapy and concurrent cisplatin or concurrent cetuximab: A surveillance, epidemiology, and end results-Medicare analysis. Cancer. 2018;124(23):4486–94.

321. Hershman DL, Wright JD, Lim E, Buono DL, Tsai WY, Neugut AI. Contraindicated use of bevacizumab and toxicity in elderly patients with cancer. J Clin Oncol. 2013;31(28):3592–9.

322. Mohile SG, Hardt M, Tew W, Owusu C, Klepin H, Gross C, *et al.* Toxicity of bevacizumab in combination with chemotherapy in older patients. Oncologist. 2013;18(4):408–14.

323. Enewold L, Sharon E, Harlan LC. Metastatic melanoma: treatment and survival in the US after the introduction of ipilimumab and vemurafenib. Oncol Res Treat. 2017;40(4):174–83.

324. Enewold L, Thomas A. Real-world Patterns of EGFR testing and treatment with erlotinib for non-small cell lung cancer in the United States. PLoS One. 2016;11(6).

325. Palazzo LL, Sheehan DF, Tramontano AC, Kong CY. Disparities and trends in genetic testing and erlotinib treatment among metastatic non-small cell lung cancer patients. Cancer Epidemiol Biomarkers Prev. 2019;28(5):926–34.

326. Maguire FB, Morris CR, Parikh-Patel A, Cress RD, Keegan THM, Li CS, *et al.* Disparities in systemic treatment use in advanced-stage non-small cell lung cancer by source of health insurance. Cancer Epidemiol Biomarkers Prev. 2019a;28(6):1059–66.

327. Thavendiranathan P, Abdel-Qadir H, Fischer HD, Liu Y, Camacho X, Amir E, *et al.* Risk-imaging mismatch in cardiac imaging practices for women receiving systemic therapy for early-stage breast cancer: a population-based cohort study. J Clin Oncol. 2018;36(30):2980–7.

328. Delate T, Won K, Carroll N, Kushi L, Hornbrook M, Bowles E, *et al.* Factors associated with first-line bevacizumab use in advanced non-squamous and non-small cell lung cancer. J Cancer Res Ther. 2014;2(1):1–8.

329. Carroll NM, Delate T, Menter A, Hornbrook MC, Kushi L, Aiello Bowles EJ, *et al.* Use of bevacizumab in community settings: toxicity profile and risk of hospitalization in patients with advanced non-small-cell lung cancer. J Oncol Pract. 2015;11(5):356–62.

330. Aarts MJ, Lemmens VEPP, Louwman MWJ, Kunst AE, Coebergh JWW. Socioeconomic status and changing inequalities in colorectal cancer? A review of the associations with risk, treatment and outcome. Eur J Cancer. 2010;46(15):2681–95.

331. Smith D, Thomson K, Bambra C, Todd A. The breast cancer paradox: A systematic review of the association between area-level deprivation and breast cancer screening uptake in Europe. Cancer Epidemiol. 2019;60:77–85.

332. Lyratzopoulos G, Abel GA, Brown CH, Rous BA, Vernon SA, Roland M, *et al.* Socio-demographic inequalities in stage of cancer diagnosis: evidence from patients with female breast, lung, colon, rectal, prostate, renal, bladder, melanoma, ovarian and endometrial cancer. Ann Oncol. 2013;24(3):843–50.

333. Hayes L, Adams J, McCallum I, Forrest L, Hidajat M, White M, *et al.* Age-related and socioeconomic inequalities in timeliness of referral and start of treatment in colorectal cancer: a population-based analysis. J Epidemiol Community Health. 2021;75(1):1–9.

334. The National Cancer Network. Evidence to March 2010 on cancer inequalities in England. [Internet]. 2010 [cited September 24th 2020]. Available from: www.ncin.org.uk/inequalities.

335. Sarfati D, Koczwara B, Jackson C. The impact of comorbidity on cancer and its treatment. CA Cancer J Clin. 2016;66(4):337–50.

336. Duma N, Idossa DW, Durani U, Frank RD, Paludo J, Westin G, *et al.* Influence of sociodemographic factors on treatment decisions in non-small-cell lung cancer. Clin Lung Cancer. 2020;21(3):e115–29.

337. Alcaraz KI, Wiedt TL, Daniels EC, Yabroff KR, Guerra CE, Wender RC. Understanding and addressing social determinants to advance cancer health equity in the United States: a blueprint for practice, research, and policy. CA Cancer J Clin. 2020;70(1):31–46.

338. Osborne CK. Steroid hormone receptors in breast cancer management. Breast Cancer Res Treat. 1998;51(3):227–38.

339. Rodgers E. Diffusion of innovations. 5th ed. New York: Somon & Schuster International; 2003.

340. Tseng CH, Chiang CJ, Tseng JS, Yang TY, Hsu KH, Chen KC, *et al.* EGFR mutation, smoking, and gender in advanced lung adenocarcinoma. Oncotarget. 2017;8(58):98384–93.

341. Lung Cancer Coalition. Molecules matter turning the science of molecular 924 diagnostics in lung cancer into a practical service for all patients [Internet]. 2019 [cited September 24th 2020]. Available from: https://www.uklcc.org.uk/wp-content/uploads/2019/01/Molecules-Matter.pdf.

342. Lee DH, Tsao MS, Kambartel KO, Isobe H, Huang MS, Barrios CH, *et al.* Molecular testing and treatment patterns for patients with advanced non-small cell lung cancer: PIvOTAL observational study. PLoS One. 2018;13(8):e0202865.

343. Illei PB, Wong W, Wu N, Chu L, Gupta R, Schulze K, *et al.* ALK Testing Trends and Patterns Among Community Practices in the United States. JCO Precis Oncol. 2018;2:1–11.

344. Cancer Research UK. Policy statement: patient access to 935 molecular diagnostics and targeted medicines in England [Internet]. 2018 [cited September 24th 2020]. Available from: https://www.cancerresearchuk.org/sites/default/files/access_to_molecular_diagnostic_tests_and_targeted_medicines_in_england_0.pdf.

345. Vogler C, Pahl J. Money, power and inequality within marriage. Sociol Rev. 1994;42(2):263–88.

346. Duncan GJ, Daly MC, McDonough P, Williams DR. Optimal indicators of socioeconomic status for health research. Am J Public Health. 2002;92(7):1151–7.

347. Piantadosi S, Byar DP, Green SB. The ecological fallacy. Am J Epidemiol. 1988;127(5):893–904.

348. Burke JF, Sussman JB, Kent DM, Hayward RA. Three simple rules to ensure reasonably credible subgroup analyses. BMJ. 2015;351:h5651.

349. Jackson D, Turner R. Power analysis for random-effects meta-analysis. Res Synth Methods. 2017;8(3):290–302.

350. Paez A. Grey literature: An important resource in systematic reviews. J Evid-Based Med. 2017;doi: 10.1111/jebm.12265.

351. Adams J, Hillier-Brown FC, Moore HJ, Lake AA, Araujo-Soares V, White M, *et al.* Searching and synthesising 'grey literature' and 'grey information' in public health: critical reflections on three case studies. Syst Rev. 2016;5(1):164.

352. Gainor JF, Shaw AT. Novel targets in non-small cell lung cancer: ROS1 and RET fusions. Oncologist. 2013;18(7):865–75.

353. Larson KL, Huang B, Chen Q, Tucker T, Schuh M, Arnold SM, *et al.* EGFR testing and erlotinib use in non-small cell lung cancer patients in Kentucky. PLoS One. 2020;15(8):e0237790.

354. Bright CJ, Lawton S, Benson S, Bomb M, Dodwell D, Henson KE, *et al.* Data resource profile: the systemic anti-cancer therapy (SACT) dataset. Int J Epidemiol. 2020;49(1):15–15l.

355. Bach PB, Guadagnoli E, Schrag D, Schussler N, Warren JL. Patient demographic and socioeconomic characteristics in the SEER-Medicare database applications and limitations. Med Care. 2002;40(8 Suppl):IV-19–25.

356. Goss E, Lopez AM, Brown CL, Wollins DS, Brawley OW, Raghavan D. American society of clinical oncology policy statement: disparities in cancer care. J Clin Oncol. 2009;27(17):2881–5.

357. Shigematsu H, Lin L, Takahashi T, Nomura M, Suzuki M, Wistuba II, *et al.* Clinical and biological features associated with epidermal growth factor receptor gene mutations in lung cancers. J Natl Cancer Inst. 2005;97(5):339–46.

358. Lauderdale DS, Goldberg J. The expanded racial and ethnic codes in the Medicare data files: their completeness of coverage and accuracy. Am J Public Health. 1996;86(5):712–6.

359. Shen C, Holguin RAP, Schaefer E, Zhou S, Belani CP, Ma PC, *et al.* Utilization and costs of epidermal growth factor receptor mutation testing and targeted therapy in Medicare patients with metastatic lung adenocarcinoma. BMC Health Serv Res. 2022;22(1):470.

360. Moser SS, Apter L, Arunachalam A, Burke T, Shalev V, Chodick G, *et al.* Real-world study of PD-L1 testing patterns and treatment distribution in patients with metastatic non-small-cell lung cancer in Israel. Immunotherapy. 2021;13(10):851–61.

361. Rivera MP, Charlot M, Durham DD, Throneburg A, Lane LM, Perera P, *et al.* Molecular biomarker and programmed death-ligand 1 expression testing in patients with advanced stage non-small cell lung cancer across North Carolina community hospitals. Chest. 2021;S0012-3692(21)00702-9.

362. Bishnoi R, Shah C, Blaes A, Bian J, Hong YR. Cardiovascular toxicity in patients treated with immunotherapy for metastatic non-small cell lung cancer: A SEER-medicare study: CVD outcomes with the use of ICI in mNSCLC. Lung Cancer. 2020;150:172–7.

363. Goulart BHL, Chennupati S, Fedorenko CR, Ramsey SD. Access to tyrosine kinase inhibitors and survival in patients with advanced EGFR+ and ALK+ positive non-small-cell lung cancer treated in the real-world. Clin Lung Cancer. 2021;S1525-7304(21)00022-X.

364. Kehl KL, Greenwald S, Chamoun NG, Manberg PJ, Schrag D. Association between first-line immune checkpoint inhibition and survival for medicare-insured patients with advanced non-small cell lung cancer. JAMA Netw Open. 2021;4(5):e2111113.

365. Li M, Diao Y, Ye J, Sun J, Jiang Y. The public health insurance coverage of novel targeted anticancer medicines in China - in vavor of whom? A retrospective analysis of the insurance claim data. Front Pharmacol. 2021;12:778940.

366. Shah M, Parmar A, Chan KKW. Socioeconomic disparity trends in diagnostic imaging, treatments, and survival for non-small cell lung cancer 2007-2016. Cancer Med. 2020 May;9(10):3407–16.

367. Lim JT, Koh JJE, Ho PJ, Liu J, Lim SH, Tan EY, *et al.* Impact of subsidy on the use of personalized medicine in breast cancer. Value Health Reg Issues. 2022;29:108–15.

368. Kaufman PA, Hurvitz SA, O'Shaughnessy J, Mason G, Yardley DA, Brufsky AM, *et al.* Baseline characteristics and first-line treatment patterns in patients with HER2-positive metastatic breast cancer in the SystHERs registry. Breast Cancer Res Treat. 2021;188(1):179–90.

369. Nguy S, Wu SP, Oh C, Gerber NK. Outcomes of HER2-positive non-metastatic breast cancer patients treated with anti-HER2 therapy without chemotherapy. Breast Cancer Res Treat. 2021;187(3):815–30.

370. Diao Y, Lin M, Xu K, Huang J, Wu X, Li M, *et al.* How government health insurance coverage of novel anti-cancer medicines benefited patients in China - a retrospective analysis of hospital clinical data. BMC Health Serv Res. 2021;21(1):856.

371. Crimini E, Repetto M, Aftimos P, Botticelli A, Marchetti P, Curigliano G. Precision medicine in breast cancer: from clinical trials to clinical practice. Cancer Treat Rev. 2021;98:102223.

372. Hudis CA. Trastuzumab - mechanism of action and use in clinical practice. N Engl J Med. 2007;357(1):39–51.

373. Slamon D, Pegram M. Rationale for trastuzumab (Herceptin) in adjuvant breast cancer trials. Semin Oncol. 2001;28(1 Suppl 3):13–9.

374. Slamon DJ, Clark GM, Wong SG, Levin WJ, Ullrich A, McGuire WL. Human breast cancer: correlation of relapse and survival with amplification of the HER-2/neu oncogene. Science. 1987;235(4785):177–82.

375. Emens LA, Esteva FJ, Beresford M, Saura C, De Laurentiis M, Kim SB, *et al.* Trastuzumab emtansine plus atezolizumab versus trastuzumab emtansine plus placebo in previously treated, HER2-positive advanced breast cancer (KATE2): a phase 2, multicentre, randomised, double-blind trial. Lancet Oncol. 2020;21(10):1283–95.

376. Cardoso F, Kyriakides S, Ohno S, Penault-Llorca F, Poortmans P, Rubio IT, *et al.* Early breast cancer: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. Ann Oncol. 2019;30(8):1194–220.

377. Gennari A, André F, Barrios CH, Cortés J, de Azambuja E, DeMichele A, *et al.* ESMO Clinical Practice Guideline for the diagnosis, staging and treatment of patients with metastatic breast cancer. Ann Oncol. 2021;32(12):1475–95.

378. Tang M, Schaffer AL, Kiely BE, Daniels B, Lee CK, Simes RJ, *et al.* Cardiac assessment in Australian patients receiving (neo)adjuvant trastuzumab for HER2-positive early breast cancer: a population-based study. Breast Cancer Res Treat. 2021;187:893–902.

379. Henson KE, Elliss-Brookes L, Coupland VH, Payne E, Vernon S, Rous B, *et al.* Data resource profile: National Cancer Registration Dataset in England. Int J Epidemiol. 2020;49(1):16–16h.

380. Public Health England. National Cancer Registration and Analysis Service (NCRAS) [Internet]. 2020 [cited August 31st 2021]. Available from: https://www.gov.uk/guidance/national-cancer-registration-and-analysis-service-ncras.

381. Public Health England. A guide to NCRAS data and its availability [Internet]. 2020 [cited November 11th 2022]. Available from: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/1073359/NCRAS_data_and_its_availability-October_2020.pdf.

382. National Cancer Registration and Analysis Service. Cancer Outcomes and Services Dataset (COSD) [Internet]. 2022 [cited November 11th 2022]. Available from: http://www.ncin.org.uk/collecting_and_using_data/data_collection/cosd#about.

383. Department for Communities and Local Government. The English Index of Multiple Deprivation (IMD) 2015 – guidance [Internet]. 2015 [cited August 28th 2021]. Available from: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/464430/English_Index_of_Multiple_Deprivation_2015_-_Guidance.pdf

384. Boyle JM, Kuryba A, Braun MS, Aggarwal A, van der Meulen J, Cowling TE, *et al.* Validity of chemotherapy information derived from routinely collected healthcare data: A national cohort study of colon cancer patients. Cancer Epidemiol. 2021;73:101971.

385. Public Health England. SACT dataset v3.0 frequently asked questions [Internet]. 2019 [cited September 7th 2021]. Available from: http://www.chemodataset.nhs.uk/frequently_asked_questions/.

386. Pickering L, Larkin J. Systemic anti-cancer therapy (SACT) dataset. Lancet Oncol. 2014;15(10):1063.

387. Public Health England. Systemic Anti-Cancer Therapy (SACT) data set technical guidance version 3.0 [Internet]. 2018 [cited November 12th 2022]. Available from: http://www.chemodataset.nhs.uk/home.

388. Public Health England. Systemic anti-cancer therapy (SACT) data set user guide [Internet]. 2018 [cited November 12th 2022]. Available from: http://www.chemodataset.nhs.uk/home.

389. NHS Digital. Systemic anti-cancer therapy (SACT) data set [Internet]. 2022 [cited November 12th 2022]. Available from: https://digital.nhs.uk/ndrs/data/data-sets/sact/sact.

390. Public Health England. Accessing PHE data through the Office for Data Release [Internet]. 2020 [cited January 5th 2022]. Available from: https://www.gov.uk/government/publications/accessing-public-health-england-data/about-the-phe-odr-and-accessing-data.

391. National Cancer Registration and Analysis Service. SACT Systemic Anti-Cancer Therapy Dataset [Internet]. 2022 [cited May 19th 2022]. Available from: http://www.chemodataset.nhs.uk/home.

392. Pathak R, Wallington M, Saunders C, Braun M, Mullamitha S, Wilson G, *et al.* Rapid analysis of outcomes using the systemic anti-cancer therapy (SACT) dataset. Clin Oncol (R Coll Radiol). 2017;29(7):e134–6.

393. Public Health England. Systemic anti-cancer therapy (SACT) - frequently asked questions (FAQs) about the SACT dataset and its role in the NICE cancer drug fund process [Internet]. 2020 [cited September 7th 2021]. Available from: http://www.chemodataset.nhs.uk/frequently_asked_questions/.

394. Royal College of Physicians. National lung cancer audit [Internet]. 2021 [cited November 12th 2022]. Available from: https://nlca.rcp.ac.uk.

395. Henson KE, Fry A, Lyratzopoulos G, Peake M, Roberts KJ, McPhail S. Sociodemographic variation in the use of chemotherapy and radiotherapy in patients with stage IV lung, oesophageal, stomach and pancreatic cancer: evidence from population-based data in England during 2013–2014. Br J Cancer. 2018;118(10):1382–90.

396. All.Can. Systemic anti-cancer therapy (SACT) data set: collecting and using real-world data to improve cancer care [Internet]. 2022 [cited November 12th 2022]. Available from: https://www.all-can.org/efficiency-hub/systemic-anti-cancer-therapy-sact-data-set-collecting-and-using-real-world-data-to-improve-cancer-care/.

397. National Cancer Registration and Analysis Service. Cancer information tools [Internet]. 2022 [cited November 12th 2022]. Available from: http://www.ncin.org.uk/cancer_information_tools/.

398. National Disease Registration Service. Welcome to cancer data [Internet]. 2022 [cited November 12th 2022]. Available from: https://www.cancerdata.nhs.uk.

399. Health Data Insight. Simulacrum [Internet]. 2022 [cited November 12th 2022]. Available from: https://simulacrum.healthdatainsight.org.uk.

400. Fraser J, Wills L, Fardus-Reid F, Irvine L, Elliss-Brookes L, Fern L, *et al.* Oral etoposide as a single agent in childhood and young adult cancer in England: still a poorly evaluated palliative treatment. Pediatr Blood Cancer. 2021;e29204.

401. Wallington M, Saxon EB, Bomb M, Smittenaar R, Wickenden M, McPhail S, *et al.* 30-day mortality after systemic anticancer treatment for breast and lung cancer in England: a population-based, observational study. Lancet Oncol. 2016;17(9):1203–16.

402. Public Health England. SACT dataset version 3.0 dataset logic document [Internet]. 2019 [cited September 7th 2021]. Available from: http://www.chemodataset.nhs.uk/frequently_asked_questions/.

403. Board R, Smittenaar R, Lawton S, Liu H, Juwa B, Chao D, *et al.* Metastatic melanoma patient outcomes since introduction of immune checkpoint inhibitors in England between 2014 and 2018. Int J Cancer. 2021;148(4):868–75.

404. Elsada A, Zalin-Miller A, Knott C, Caravotas L. A registry study of relapsed or refractory multiple myeloma pre-exposed to three or more prior therapies including a proteasome

inhibitor, an immunomodulatory agent and CD38-targeted monoclonal antibody therapy in England. EJHaem. 2021;2(3):493-97.

405. Jones GS, McKeever TM, Hubbard RB, Khakwani A, Baldwin DR. Factors influencing treatment selection and 30-day mortality after chemotherapy for people with small-cell lung cancer: an analysis of national audit data. Eur J Cancer. 2018;103:176–83.

406. Mintz HP. Can routinely collected data be used to inform randomised controlled trial outcomes in oncology? [PhD Thesis]. University of Warwick; 2019.

407. Hounsome L, Eyre TA, Ireland R, Hodson A, Walewska R, Ardeshna K, *et al.* Diffuse large B cell lymphoma (DLBCL) in patients older than 65 years: analysis of 3 year real world data of practice patterns and outcomes in England. Br J Cancer. 2022;126(1):134–43.

408. Rahman S, Thomas B, Maynard N, Park MH, Wahedally M, Trudgill N, *et al.* Impact of postoperative chemotherapy on survival for oesophagogastric adenocarcinoma after preoperative chemotherapy and surgery. Br J Surg. 2022;109(2):227–36.

409. Boyle JM, Kuryba A, Cowling TE, van der Meulen J, Fearnhead NS, Walker K, *et al.* Survival outcomes associated with completion of adjuvant oxaliplatin-based chemotherapy for stage III colon cancer: A national population-based study. Int J Cancer. 2022;150(2):335–46.

410. Boyle JM, Cowling TE, Kuryba A, Fearnhead NS, van der Meulen J, Braun MS, *et al.* Development and validation of a coding framework to identify severe acute toxicity from systemic anti-cancer therapy using hospital administrative data. Cancer Epidemiol. 2022;77:102096.

411. McDonald L, Sammon C, Carroll R, Harish A, Shankar A, Tyas D, *et al.* Consistency of recording of chemotherapy cycles in the National Cancer Registration and Analysis Service Systemic Anti-Cancer Therapy database and the hospital episode statistics admitted patient care database. Future Oncol. 2020;16(3):4455–60.

412. Horvat P, Gray CM, Lambova A, Christian JB, Lasiter L, Stewart M, *et al.* Comparing findings from a friends of cancer research exploratory analysis of real-world end points with the cancer analysis system in England. JCO Clin Cancer Inform. 2021;5:1155–68.

413. Khoja L, McGurk A, O'Hara C, Chow S, Hasan J. Mortality within 30 days following systemic anti-cancer therapy, a review of all cases over a 4 year period in a tertiary cancer centre. Eur J Cancer. 2015 Jan;51(2):233–40.

414. O'Brien MER, Borthwick A, Rigg A, Leary A, Assersohn L, Last K, *et al.* Mortality within 30 days of chemotherapy: a clinical governance benchmarking issue for oncology patients. Br J Cancer. 2006;95(12):1632–6.

415. Wadd N, Peedell C, Polwart C. Real-world assessment of cancer drugs using local data uploaded to the Systemic Anti-Cancer Therapy Dataset in England. Clin Oncol (R Coll Radiol). 2022;S0936-6555(22)00221-7.

416. Sweeting MJ, Oliver-Williams C, Teece L, Welch CA, de Belder MA, Coles B, *et al.* Data resource profile: the Virtual Cardio-Oncology Research Initiative (VICORI) linking national English cancer registration and cardiovascular audits. Int J Epidemiol. 2022;50(6):1768-1779.

417. Downing A, Hall P, Birch R, Lemmon E, Affleck P, Rossington H, *et al.* Data resource profile: the COloRECTal cancer data repository (CORECT-R). Int J Epidemiol. 2021 ;50(5):1418-1418k.

418. Taylor J, Wright P, Rossington H, Mara J, Glover A, West N, *et al.* Regional multidisciplinary team intervention programme to improve colorectal cancer outcomes: study protocol for the Yorkshire Cancer Research Bowel Cancer Improvement Programme (YCR BCIP). BMJ Open. 2019;9(11):e030618.

419. Taylor JC, Swinson D, Seligmann JF, Birch RJ, Dewdney A, Brown V, *et al.* Addressing the variation in adjuvant chemotherapy treatment for colorectal cancer: can a regional intervention promote national change? Int J Cancer. 2021;148(4):845–56.

420. Gathani T, Reeves G, Broggio J, Barnes I. Ethnicity and the tumour characteristics of invasive breast cancer in over 116,500 women in England. Br J Cancer. 2021;125(4):611–7.

421. National Disease Registration Service. Cancer Outcomes and Services Dataset (COSD) user guide v9.1.3 [Internet]. 2021 [cited December 12th 2022]. Available from: http://www.ncin.org.uk/collecting_and_using_data/data_collection/cosd_downloads_v9 .

422. Sterne JAC, White IR, Carlin JB, Spratt M, Royston P, Kenward MG, *et al.* Multiple imputation for missing data in epidemiological and clinical research: potential and pitfalls. BMJ. 2009;338:b2393.

423. Jakobsen JC, Gluud C, Wetterslev J, Winkel P. When and how should multiple imputation be used for handling missing data in randomised clinical trials - a practical guide with flowcharts. BMC Med Res Methodol. 2017;17(1):162.

424. Cancer Research UK. Find a clinical trial [Internet]. 2022 [cited June 23rd 2022]. Available from: https://www.cancerresearchuk.org/about-cancer/find-a-clinical-trial.

425. National Institute of Health. Home - ClinicalTrials.gov [Internet]. 2022 [cited June 23rd 2022]. Available from: https://www.clinicaltrials.gov/.

426. National Institute for Health and Care Excellence. Overview guidance on the use of trastuzumab for the treatment of advanced breast cancer [Internet]. 2002 [cited September 17th 2021]. Available from: https://www.nice.org.uk/guidance/ta34.

427. Electronics Medicines Compendium. Summary of product characteristics [Internet]. 2022 [cited March 24th 2022]. Available from: https://www.medicines.org.uk/emc#gref.

428. European Medicines Agency. European Medicines Agency [Internet]. 2022 [cited March 24th 2022]. Available from: https://www.ema.europa.eu/en.

429. Northern Cancer Alliance. Northern Cancer Alliance protocols [Internet]. 2020 [cited March 24th 2022]. Available from: https://northerncanceralliance.nhs.uk/.

430. European Society for Medical Oncology. ESMO guidelines [Internet]. 2022 [cited March 22nd 2022]. Available from: https://www.esmo.org/guidelines.

431. Department for Communities and Local Government. English indices of deprivation 2010: overall [Internet]. 2010 [cited August 28th 2021]. Available from: https://www.gov.uk/government/statistics/english-indices-of-deprivation-2010.

432. Department for Communities and Local Government. File 2 domains of deprivation [Internet]. 2015 [cited August 28th 2021]. Available from: https://www.gov.uk/government/statistics/english-indices-of-deprivation-2015.

433. National Cancer Registration and Analysis Service. Collecting and using data. [Internet] 2023 [cited January 27th 2023]. Available from: http://ncin.org.uk/collecting_and_using_data/.

434. Bibby P, Brindley P. The 2011 rural-urban classification for small area geographies: a user guide and frequently asked questions (v1.0) [Internet]. Office for National Statistics; 2013 [cited May 19th 2022]. Available from: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/239478/RUC11user_guide_28_Aug.pdf.

435. Office for National Statistics. England: Detailed information on the administrative structure within England. [Internet]. 2021 [cited August 30th 2021]. Available from: https://www.ons.gov.uk/methodology/geography/ukgeographies/administrativegeography/england.

436. Gómez HL, Castañeda CA, Vigil CE, Vidaurre T, Velarde RG, Cruz WR, *et al.* Prognostic effect of hormone receptor status in early HER2 positive breast cancer patients. Hematol Oncol Stem Cell Ther. 2010;3(3):109–15.

437. Charlson ME, Pompei P, Ales KL, MacKenzie CR. A new method of classifying prognostic comorbidity in longitudinal studies: development and validation. J Chronic Dis. 1987;40(5):373–83.

438. Schneeweiss S, Maclure M. Use of comorbidity scores for control of confounding in studies using administrative databases. Int J Epidemiol. 2000;29(5):891–8.

439. Quan H, Sundararajan V, Halfon P, Fong A, Burnand B, Luthi JC, *et al.* Coding algorithms for defining comorbidities in ICD-9-CM and ICD-10 administrative data. Med Care. 2005;43(11):1130–9.

440. Hosmer DW, Lemeshow S. Applied logistic regression. New York: Wiley; 2013.

441. Sankar P, Cho MK, Condit CM, Hunt LM, Koenig B, Marshall P, *et al.* Genetic research and health disparities. JAMA. 2004;291(24):2985–9.

442. Veugelers P, Yip A. Socioeconomic disparities in health care use: Does universal coverage reduce inequalities in health? J Epidemiol Community Health. 2003;57(6):424–8.

443. McKeage K, Perry CM. Trastuzumab: a review of its use in the treatment of metastatic breast cancer overexpressing HER2. Drugs. 2002;62(1):209–43.

444. McDermott AM, Wall DM, Waters PS, Cheung S, Sibbering M, Horgan K, *et al.* Surgeon and breast unit volume-outcome relationships in breast cancer surgery and treatment. Ann Surg. 2013;258(5):808–13; discussion 813-814.

445. Richards M, Sainsbury R, Kerr D. Inequalities in breast cancer care and outcome. Br J Cancer. 1997;76(5):634–8.

446. Henley NC, Hole DJ, Kesson E, Burns HJG, George WD, Cooke TG. Does deprivation affect breast cancer management? Br J Cancer. 2005;92(4):631–3.

447. Lian W, Fu F, Lin Y, Lu M, Chen B, Yang P, *et al.* The impact of young age for prognosis by subtype in women with early breast cancer. Sci Rep. 2017;7(1):11625.

448. Lima SM, Kehm RD, Swett K, Gonsalves L, Terry MB. Trends in parity and breast cancer incidence in US women younger than 40 years from 1935 to 2015. JAMA Netw Open. 2020;3(3):e200929.

449. International Agency for Research on Cancer. Cancer incidence in five continents vol. XI number 166 volume XI [Internet]. 2021 [cited January 27th 2023]. Available from: https://publications.iarc.fr/Book-And-Report-Series/Iarc-Scientific-Publications/Cancer-Incidence-In-Five-Continents%C2%A0Volume-XI-2021.

450. Chiu C, Morrell S, Page A, Rickard M, Brassil A, Taylor R. Population-based mammography screening and breast cancer incidence in New South Wales, Australia. Cancer Causes Control. 2006;17(2):153–60.

451. Martins T, Abel G, Ukoumunne OC, Price S, Lyratzopoulos G, Chinegwundoh F, *et al.* Assessing ethnic inequalities in diagnostic interval of common cancers: a population-based UK cohort Study. Cancers. 2022;14(13):3085.

452. Saunders C, Abel G, Lyratzopoulos G. Inequalities in reported cancer patient experience by socio-demographic characteristic and cancer site: evidence from respondents to the English cancer patient experience survey. Eur J Cancer Care. 2015;24(1):85–98.

453. Hayanga B, Stafford M, Bécares L. Ethnic inequalities in healthcare use and care quality among people with multiple long-term health conditions living in the United Kingdom:

a systematic review and narrative synthesis. Int J Environ Res Public Health. 2021;18(23):12599.

454. Wong EC, Collins RL, Breslau J, Burnam MA, Cefalu M, Roth EA. Differential association of stigma with perceived need and mental health service use. J Nerv Ment Dis. 2018;206(6):461–8.

455. Chambers P, Man KKC, Lui VWY, Mpima S, Nasuti P, Forster MD, *et al.* Understanding molecular testing uptake across tumor types in eight countries: results from a multinational cross-sectional survey. JCO Oncol Pract. 2020;16(8):e770–8.

456. Zubrod CG, Schneiderman M, Frei E, Brindley C, Lennard Gold G, Shnider B, *et al.* Appraisal of methods for the study of chemotherapy of cancer in man: comparative therapeutic trial of nitrogen mustard and triethylene thiophosphoramide. J Chronic Dis. 1960;11(1):7–33.

457. Karnofsky DA, Abelmann WH, Craver LF, Burchenal JH. The use of the nitrogen mustards in the palliative treatment of carcinoma. With particular reference to bronchogenic carcinoma. Cancer. 1948;1(4):634–56.

458. Broderick JE, May M, Schwartz JE, Li M, Mejia A, Nocera L, *et al.* Patient reported outcomes can improve performance status assessment: a pilot study. J Patient-Rep Outcomes. 2019;3(1):41.

459. Landry LG, Ali N, Williams DR, Rehm HL, Bonham VL. Lack of diversity in genomic databases is a barrier to translating precision medicine research into practice. Health Aff. 2018 y;37(5):780–5.

460. Chakravarthy R, Stallings SC, Williams M, Hollister M, Davidson M, Canedo J, *et al.* Factors influencing precision medicine knowledge and attitudes. PLoS One. 2020;15(11):e0234833.

461. Yu M, Feuer EJ, Cronin KA, Caporaso NE. Use of multiple imputation to correct for bias in lung cancer incidence trends by histologic subtype. Cancer Epidemiol Biomark Prev l. 2014;23(8):1546–58.

462. Surveillance, Epidemiology, and End Results Program. ICD-0-3 SEER site/histology validation list [Internet]. 2021 [cited August 27th 2021]. Available from: https://seer.cancer.gov/icd-o-3/sitetype.icdo3.20210607.pdf.

463. National Cancer Institute. Morphology & grade SEER training [Internet]. 2021 [cited August 27th 2021]. Available from: https://training.seer.cancer.gov/lung/abstract-code-stage/morphology.html.

464. Ries LAG, Young JL, Keel GE, Eisner MP, Lin YD, Horner MJD (eds). SEER survival monograph: cancer survival among adults. U.S. SEER program, 1988-2001 patient and tumour characteristics. National Cancer Institute, SEER Program, NIH pub no.07-6215, MD; 2007.

465. Lewis DR, Check DP, Caporaso NE, Travis WD, Devesa SS. US lung cancer trends by histologic type. Cancer. 2014;120(18):2883–92.

466. Forrest LF. Intervention-generated inequalities in lung cancer care [PhD Thesis] Newcastle University; 2014.

467. Postmus PE, Kerr KM, Oudkerk M, Senan S, Waller DA, Vansteenkiste J, *et al.* Early and locally advanced non-small-cell lung cancer (NSCLC): ESMO clinical practice guidelines for diagnosis, treatment and follow-up. Ann Oncol. 2017;28(suppl_4):iv1–21.

468. Lindeman NI, Cagle PT, Aisner DL, Arcila ME, Beasley MB, Bernicker EH, *et al.* Updated molecular testing guideline for the selection of lung cancer patients for treatment with targeted tyrosine kinase inhibitors: guideline from the College of American pathologists, the international association for the study of lung cancer, and the association for molecular pathology. Arch Pathol Lab Med. 2018;142(3):321–46.

469. Lin Y, Mak KS. Racial and socioeconomic disparities in the use of stereotactic body radiotherapy for treating non-small cell lung cancer: a narrative review. J Thorac Dis. 2021;13(6):3764–71.

470. Kerr KM, Bibeau F, Thunnissen E, Botling J, Ryška A, Wolf J, *et al*. The evolving landscape of biomarker testing for non-small cell lung cancer in Europe. Lung Cancer. 2021;154:161–75.

471. Hsu WH, Yang JCH, Mok TS, Loong HH. Overview of current systemic management of EGFR-mutant NSCLC. Ann Oncol. 2018;29(suppl_1):i3–9.

472. Bennett V, Davies E, Jack R, Moller H. Histological subtype of lung cancer in relation to socio-economic deprivation in south east England. BMC Cancer. 2008;8:139.

473. Zhao W, Jiang W, Wang H, He J, Su C, Yu Q. Impact of smoking history on response to immunotherapy in non-small-cell lung cancer: a systematic review and meta-analysis. Front Oncol. 2021;11:703143.

474. Suh WN, Kong KA, Han Y, Kim SJ, Lee SH, Ryu YJ, *et al.* Risk factors associated with treatment refusal in lung cancer. Thorac Cancer. 2017;8(5):443–50.

475. Frank JW, Matsunaga E. National monitoring systems for health inequalities by socioeconomic status - an OECD snapshot. Crit Public Health. 2020;32(4):doi: 10.1080/09581596.2020.1862761:580–7.

476. NHS Digital. NDRS somatic molecular dataset [Internet]. 2022 [cited November 15th 2022]. Available from: https://digital.nhs.uk/services/data-access-request-service-dars/dars-products-and-services/data-set-catalogue/ndrs-somatic-molecular-data-set.

477. Jack RH, Linklater KM, Hofman D, Fitzpatrick J, Møller H. Ethnicity coding in a regional cancer registry and in hospital episode statistics. BMC Public Health. 2006;10(6):281.

478. Booth CM, Tannock IF. Randomised controlled trials and population-based observational research: partners in the evolution of medical evidence. Br J Cancer. 2014;110(3):551–5.

479. Gibelin C, Couraud S. Somatic alterations in lung cancer: do environmental factors matter? Lung Cancer. 2016;100:45–52.

480. Erhunmwunsee L, Wing SE, Shen J, Hu H, Sosa E, Lopez LN, *et al.* The association between polluted neighborhoods and TP53-mutated non-small cell lung cancer. Cancer Epidemiol Biomarkers Prev. 2021;30(8):1498–1505.

481. Mogi A, Kuwano H. TP53 mutations in nonsmall cell lung cancer. J Biomed Biotechnol. 2011;2011:583929.

482. Marmot M. Health equity, cancer, and social determinants of health. Lancet Glob Health. 2018;6:S29.

483. Stitzenberg KB, Sigurdson ER, Egleston BL, Starkey RB, Meropol NJ. Centralization of cancer surgery: implications for patient access to optimal care. J Clin Oncol. 2009;27(28):4671–8.

484. Baldwin DR. Socioeconomic position and delays in lung cancer diagnosis: should we target the more deprived? Thorax. 2017;72(5):393–5.

485. Lyratzopoulos G, Neal RD, Barbiere JM, Rubin GP, Abel GA. Variation in number of general practitioner consultations before hospital referral for cancer: findings from the 2010 national cancer patient experience survey in England. Lancet Oncol. 2012;13(4):353–65.

486. Bradley SH, Barclay M, Cornwell B, Abel GA, Callister ME, Gomez-Cano M, *et al.* Associations between general practice characteristics and chest X-ray rate: an observational study. Br J Gen Pract. 2021;72(714):e34–42.

487. Abel GA, Shelton J, Johnson S, Elliss-Brookes L, Lyratzopoulos G. Cancer-specific variation in emergency presentation by sex, age and deprivation across 27 common and rarer cancers. Br J Cancer. 2015;112 Suppl 1:S129-136.

488. Macmillan. The age old excuse: the under treatment of older cancer patients [Internet]. 2021 [cited February 28th 2022]. Available from: https://www.macmillan.org.uk/documents/getinvolved/campaigns/ageoldexcuse/ageoldexcusereport-macmillancancersupport.pdf

489. Smith G, Lopez-Olivo M, Advani P, Ning M, Geng Y, Giordano S, *et al.* Financial burdens of cancer treatment: a systematic review of risk factors and outcomes. J Natl Compr Cancer Netw. 2019;17(10):1184-1192.

490. Weiner S. 'I can't afford that!': dilemmas in the care of the uninsured and underinsured. J Gen Intern Med. 2001;16(6):412–8.

491. MacLeod S, Musich S, Hawkins K, Armstrong DG. The growing need for resources to help older adults manage their financial and healthcare choices. BMC Geriatr. 2017;17(1):84.

492. Carrera PM, Kantarjian HM, Blinder VS. The financial burden and distress of patients with cancer: understanding and stepping-up action on the financial toxicity of cancer treatment. CA Cancer J Clin. 2018;68(2):153-165.

493. Macmillan. Hundreds of thousands of cancer patients simply 'can't afford life' as the cost-of-living crisis risks patient's health [Internet]. 2022 [cited December 21st 2022]. Available from: https://medium.com/macmillan-press-releases-and-

statements/hundreds-of-thousands-of-cancer-patients-simply-cant-afford-life-as-the-cost-of-living-crisis-2d75712a5a41.

494. Limb M. Cancer patients' health is at increased risk from cost of living crisis, charity warns. BMJ. 2022;377:o1103.

495. The Lancet Oncology. Cancer, health-care backlogs, and the cost-of-living crisis. Lancet Oncol. 2022;23(6):691.

496. Dowsett M, Hanby AM, Laing R, Walker R. HER2 testing in the UK: consensus from a national consultation. J Clin Pathol. 2007;60(6):685–9.

497. Allemani C, Matsuda T, Di Carlo V, Harewood R, Matz M, Nikšić M, *et al.* Global surveillance of trends in cancer survival 2000-14 (CONCORD-3): analysis of individual records for 37,513025 patients diagnosed with one of 18 cancers from 322 population-based registries in 71 countries. Lancet. 2018;391(10125):1023–75.

498. De Angelis R, Sant M, Coleman MP, Francisci S, Baili P, Pierannunzio D, *et al.* Cancer survival in Europe 1999-2007 by country and age: results of EUROCARE-5 - a population-based study. Lancet Oncol. 2014;15(1):23–34.

499. Larsen MD, Cars T, Hallas J. A mini review of the use of hospital-based databases in observational inpatient studies of drugs. Basic Clin Pharmacol Toxicol. 2013;112(1):13–8.

500. Higgins J, Altman D, Sterne JAC. Chapter 8: assessing risk of bias in included studies. In: Higgins JPT and Green S, editors. Cochrane handbook for systematic reviews of interventions (version 510 [updated March 2011]). Chichester: John Wiley & Sons, Ltd; 2011.

501. Lichtensztajn DY, Giddings BM, Morris CR, Parikh-Patel A, Kizer KW. Comorbidity index in central cancer registries: the value of hospital discharge data. Clin Epidemiol. 2017;9:601–9.

502. Lüchtenborg M, Morris EJA, Tataru D, Coupland VH, Smith A, Milne RL, *et al.* Investigation of the international comparability of population-based routine hospital data set derived comorbidity scores for patients with lung cancer. Thorax. 2018;73(4):339–49.

503. Fowler H, Belot A, Ellis L, Maringe C, Luque-Fernandez MA, Njagi EN, *et al.* Comorbidity prevalence among cancer patients: a population-based cohort study of four cancers. BMC Cancer. 2020;20(1):2.

504. Louwman WJ, Aarts MJ, Houterman S, van Lenthe FJ, Coebergh JWW, Janssen-Heijnen MLG. A 50% higher prevalence of life-shortening chronic conditions among cancer patients with low socioeconomic status. Br J Cancer. 2010;103(11):1742–8.

505. Stirland LE, González-Saavedra L, Mullin DS, Ritchie CW, Muniz-Terrera G, Russ TC. Measuring multimorbidity beyond counting diseases: systematic review of community and population studies and guide to index choice. BMJ. 2020;368:m160.

506. Purushotham A, Bains S, Lewison G, Szmukler G, Sullivan R. Cancer and mental health - a clinical and research unmet need. Ann Oncol. 2013;24(9):2274–8.

507. Mental Health Foundation Scotland. Supporting the emotional and mental health needs of people with cancer [Internet]. 2018 [cited March 2nd 2022]. Available from: https://www.mentalhealth.org.uk/sites/default/files/mh-needs-people-with-cancer.pdf.

508. Moscrop A, Ziebland S, Bloch G, Iraola JR. If social determinants of health are so important, shouldn't we ask patients about them? BMJ. 2020;371:m4150.

509. Kelly S. Individual socioeconomic status is important but hard to measure. BMJ. 2021;372:n174.

510. Moss JL, Johnson NJ, Yu M, Altekruse SF, Cronin KA. Comparisons of individual and area-level socioeconomic status as proxies for individual-level measures: evidence from the mortality disparities in American communities study. Popul Health Metr. 2021;19(1): doi.org/10.1186/s12963-020-00244-x.

511. Bryere J, Pornet C, Copin N, Launay L, Gusto G, Grosclaude P, *et al.* Assessment of the ecological bias of seven aggregate social deprivation indices. BMC Public Health. 2017;17(1):86.

512. Ingleby FC, Belot A, Atherton I, Baker M, Elliss-Brookes L, Woods LM. Assessment of the concordance between individual-level and area-level measures of socio-economic deprivation in a cancer patient cohort in England and Wales. BMJ Open. 2020;10(11):e041714.

513. Lokar K, Zagar T, Zadnik V. Estimation of the ecological fallacy in the geographical analysis of the association of socio-economic deprivation and cancer incidence. Int J Environ Res Public Health. 2019;16(3):296.

514. Buajitti E, Chiodo S, Rosella LC. Agreement between area- and individual-level income measures in a population-based cohort: Implications for population health research. SSM Popul Health. 2020;10:100553.

515. Nardini C, Osmani V, Cormio PG, Frosini A, Turrini M, Lionis C, *et al.* The evolution of personalized healthcare and the pivotal role of European regions in its implementation. Pers Med. 2021;18(3):283–94.

516. Kickbusch I, Pelikan JM, Apfel F, Tsouros AD. The World Health Organisation, regional office for Europe 2013 [Internet]. 2013 [cited May 13th 2022]. Available from: https://apps.who.int/iris/bitstream/handle/10665/128703/e96854.pdf.

517. Martin LT, Schonlau M, Haas A, Rosenfeld L, Derose KP, Buka SL, *et al*. Patient activation and advocacy: which literacy skills matter most? J Health Commun. 2011;16(Suppl 3):177–190.

518. EGFR Resisters. EGFR Resisters [Internet]. 2022 [cited April 14th 2022]. Available from: https://egfrcancer.org.

519. ALK Positive UK. ALK Positive UK [Internet]. 2022 [cited April 14th 2022]. Available from: https://www.alkpositive.org.uk.

520. Breast Cancer Now. Breast Cancer now: the research and care charity [Internet]. 2022 [cited May 13th 2022]. Available from: https://breastcancernow.org/get-involved/other-ways-give/natasha-hope?gclid=EAIaIQobChMIrO_R473c9wIVkevtCh18xghSEAAYASAAEgLjW_D_BwE.

521. Public Health England. Trust-level 30-day mortality after systemic anticancer treatment for breast and lung cancer in England Companion report [Internet]. 2016 [cited April 14th 2022]. Available from: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/549299/30_day_mortality_after_sact.pdf.

522. NHS England, NHS Improvement clinically-led review of NHS cancer standards [Internet]. 2022 [cited May 13th 2022]. Available from: https://www.england.nhs.uk/wp-content/uploads/2022/03/B1320-clinically-led-review-of-nhs-cancer-standards-models-of-care-and-measurement_090322.pdf.

523. Macmillan Cancer Support. 2011-12 National cancer patient experience survey - league table and tumour group variations [Internet]. 2012 [cited June 25th 2022]. Available from: https://www.macmillan.org.uk/documents/aboutus/commissioners/patientexperiencesurvey2012.pdf.

524. Jamtvedt G, Flottorp S, Ivers N. Audit and feedback as a quality strategy. Improving healthcare quality in Europe: characteristics, effectiveness and implementation of different strategies. European observatory on health systems and policies [Internet]. 2019 [cited December 19th 2022]. Available from: https://www.ncbi.nlm.nih.gov/books/NBK549284/.

525. Ivers N, Jamtvedt G, Flottorp S, Young JM, Odgaard-Jensen J, French SD, *et al.* Audit and feedback: effects on professional practice and healthcare outcomes. Cochrane Database Syst Rev. 2012;(6):CD000259.

526. Hendryx M, Luo J. Increased cancer screening for low-income adults under the affordable care act medicaid expansion. Med Care. 2018;56(11):944–9.

527. Gao L, Joseph J, Santoro-Levy M, Multz AS, Gotlieb VK. Utilization of pharmaceutical patient and prescription assistance programs via a pharmacy department patient assistance program for indigent cancer patients. Hosp Pharm. 2016;51(7):572–6.

528. Fisher R, Allen L, Malhotra AM, Alderwick H. Tackling the inverse care law Analysis of policies to improve general practice in deprived areas since 1990. The Health Foundation, London [Internet]. 2022 [cited May 16th 2022]. Available from: https://www.health.org.uk/sites/default/files/upload/publications/2022/Tackling%20the%20inverse%20care%20law.pdf.

529. Coupe N, Cotterill S, Peters S. Tailoring lifestyle interventions to low socio-economic populations: a qualitative study. BMC Public Health. 2018;18(1):967.

530. Campbell DJT, Manns BJ, Leblanc P, Hemmelgarn BR, Sanmartin C, King-Shier K. Finding resiliency in the face of financial barriers. Medicine (Baltimore). 2016;95(49):e5561.

531. Davis TC, Bocchini JA, Fredrickson D, Arnold C, Mayeaux EJ, Murphy PW, *et al.* Parent comprehension of polio vaccine information pamphlets. Pediatrics. 1996;97(6 Pt 1):804–10.

532. Ockene IS, Tellez TL, Rosal MC, Reed GW, Mordes J, Merriam PA, *et al.* Outcomes of a Latino community-based intervention for the prevention of diabetes: the Lawrence Latino diabetes prevention project. Am J Public Health. 2012;102(2):336–42.

533. The Kings Fund. Poverty and the health and care system: the role of data and partnership in bringing change [Internet]. 2022 [cited December 22nd 2022]. Available from: https://www.kingsfund.org.uk/publications/poverty-health-care-system-data-partnership.

534. Dharamsi S, Espinoza N, Cramer C, Amin M, Bainbridge L, Poole G. Nurturing social responsibility through community service-learning: lessons learned from a pilot project. Med Teach. 2010;32(11):905–11.

535. Liddy C, Rowan MS, Afkham A, Maranger J, Keely E. Building access to specialist care through e-consultation. Open Med. 2013;7(1):e1-8.

536. The Catalyst. Citizens advice Southampton: validating an existing video hub solution providing advice to vulnerable people [Internet]. 2022 [cited February 20th 2022]. Available from: https://ochre.thecatalyst.org.uk/initiatives/199730-citizens-advice-southampton-validating-an-existing-video-hub-solution-providing-advice-to-vulnerable-people.

537. Nelson HD, Cantor A, Wagner J, Jungbauer R, Fu R, Kondo K, *et al.* Effectiveness of patient navigation to increase cancer screening in populations adversely affected by health disparities: a meta-analysis. J Gen Intern Med. 2020;35(10):3026–35.

538. Sung JF, Blumenthal DS, Coates RJ, Williams JE, Alema-Mensah E, Liff JM. Effect of a cancer screening intervention conducted by lay health workers among inner-city women. Am J Prev Med. 1997;13(1):51–7.

539. Dietrich AJ, Tobin JN, Cassells A, Robinson CM, Greene MA, Sox CH, *et al.* Telephone care management to improve cancer screening among low-income women: a randomized, controlled trial. Ann Intern Med. 2006;144(8):563–71.

540. Dranitsaris G, Truter I, Lubbe MS, Amir E, Evans W. Advances in cancer therapeutics and patient access to new drugs. Pharmacoeconomics. 2011;29(3):213–24.

541. Dranitsaris G, Truter I, Lubbe MS, Sriramanakoppa NN, Mendonca VM, Mahagaonkar SB. Improving patient access to cancer drugs in India: using economic modeling to estimate a more affordable drug cost based on measures of societal value. Int J Technol Assess Health Care. 2011;27(1):23–30.

542. Steptoe A, Perkins-Porras L, McKay C, Rink E, Hilton S, Cappuccio FP. Behavioural counselling to increase consumption of fruit and vegetables in low income adults: randomised trial. BMJ. 2003;326(7394):855.

543. Lawler M, Selby P, Aapro MS, Duffy S. Ageism in cancer care. BMJ. 2014;348:g1614.

544. Ginsburg GS, Phillips KA. Precision medicine: from science to value. Health Aff Proj Hope. 2018;37(5):694–701.

545. Martin NA, Tepper JE, Giri VN, Stinchcombe TE, Cheng HH, Javle MM, *et al.* Adopting consensus terms for testing in precision medicine. JCO Precis Oncol. 2021;5:PO.21.00027.

546. Darin-Mattsson A, Fors S, Kåreholt I. Different indicators of socioeconomic status and their relative importance as determinants of health in old age. Int J Equity Health. 2017;16(1):173.

547. Macintyre S, McKay L, Der G, Hiscock R. Socio-economic position and health: what you observe depends on how you measure it. J Public Health. 2003;25(4):288–94.

548. Osborn AF. Assessing the socio-economic status of families. Sociology. 1987;21(3):429–48.

549. Office for National Statistics. One in eight young people without degrees work in graduate jobs [Internet]. 2018 [cited August 19th 2020]. Available from: https://www.ons.gov.uk/employmentandlabourmarket/peopleinwork/employmentande

mployeetypes/articles/oneineightyoungpeoplewithoutdegreesworkingraduatejobs/2018-09-18.

550. Pickett KE, Pearl M. Multilevel analyses of neighbourhood socioeconomic context and health outcomes: a critical review. J Epidemiol Community Health. 2001;55(2):111–22.

551. Tucker TC, Durbin EB, McDowell JK, Huang B. Unlocking the potential of population-based cancer registries. Cancer. 2019;125(21):3729–37.

552. Mathur R, Bhaskaran K, Chaturvedi N, Leon DA, vanStaa T, Grundy E, *et al.* Completeness and usability of ethnicity data in UK-based primary care and hospital databases. J Public Health. 2014;36(4):684–92.

553. Heins MJ, de Ligt KM, Verloop J, Siesling S, Korevaar JC. Opportunities and obstacles in linking large health care registries: the primary secondary cancer care registry - breast cancer. BMC Med Res Methodol. 2022;22(1):124.

554. NHS Digital. Summary Care Records (SCR) [Internet]. 2022 [cited November 15th 2022]. Available from: https://digital.nhs.uk/services/summary-care-records-scr.

555. Castellanos EH, Orlando A, Ma X, Parikh RB, O'Connell G, Meropol NJ, *et al.* Evaluating the impact of oncology care model reporting requirements on biomarker testing and treatment. JCO Oncol Pract. 2020;16(10):e1216–21.

556. Pennell NA, Arcila ME, Gandara DR, West H. Biomarker testing for patients with advanced non-small cell lung cancer: real-world issues and tough choices. Am Soc Clin Oncol Edu Book. 2019;39:531–542.

557. Great Manchester Care Record. Great Manchester Care Record we are better together [Internet]. 2022 [cited April 13th 2022]. Available from: https://gmwearebettertogether.com/.

558. Great North Care Record. Great North Care Record [Internet]. 2022 [cited April 13th 2022]. Available from: https://www.greatnorthcarerecord.org.uk/.