Accounting for Capacity Constraints in Economic Evaluations of Precision Medicine

A thesis submitted to The University of Manchester for the degree of Doctor of Philosophy in the Faculty of Biology, Medicine and Health

2020

Stuart J Wright

School of Health Sciences
Division of Population Health, Health Services Research & Primary Care
Content
List of Tables .................................................................................................................. 5
List of Figures .................................................................................................................. 6
List of Appendices ......................................................................................................... 7
List of Abbreviations ..................................................................................................... 8
Glossary .......................................................................................................................... 10
Abstract .......................................................................................................................... 13
Declaration ...................................................................................................................... 15
Copyright Statement ...................................................................................................... 16
Rationale for Journal Format Submission ....................................................................... 17
Publications ..................................................................................................................... 18
Acknowledgements ........................................................................................................ 20
The Author ..................................................................................................................... 21
Chapter 1 ....................................................................................................................... 22
Introduction .................................................................................................................... 22
  1.1 Thesis Summary ...................................................................................................... 22
  1.2 Precision Medicine .............................................................................................. 22
  1.3 Opportunity Cost .................................................................................................. 24
  1.4 The Role of Economic Evaluation ....................................................................... 25
  1.5 Generating Economic Evidence ......................................................................... 27
  1.6 The Cost-Effectiveness Threshold ...................................................................... 29
  1.7 Assumptions of Decision-Analytic Model-Based Cost-Effectiveness Analysis ...... 30
  1.8 Health System Capacity Constraints .................................................................. 32
  1.9 The Importance of Understanding the Impact of Constraints in Health System Capacity to the Provision of Precision Medicine .................................................. 32
  1.10 Potential Implications of Capacity Constraints for the Economic Evaluation of Precision Medicine ................................................................. 33
  1.11 Research Questions ............................................................................................ 34
  1.13 Thesis Outline .................................................................................................... 35
Chapter 2 ...................................................................................................................... 37
Accounting for Capacity Constraints in Economic Evaluations of Precision Medicine: a Systematic Review ............................................................................................ 37
  2.1 Chapter 2 Overview .............................................................................................. 37
  2.2 Abstract ................................................................................................................ 38
  2.3 Introduction .......................................................................................................... 39
  2.4 Methods ............................................................................................................... 40
  2.5 Results ................................................................................................................ 45
  2.6 Discussion ............................................................................................................ 60
  2.7 Conclusions ........................................................................................................ 64
  2.8 Summary of Chapter 2 ...................................................................................... 65
6.7 Summary of Chapter 6.............................................................................................................188
Chapter 7 ........................................................................................................................................189
Quantifying the Value of Investments to Reduce the Impact of Capacity Constraints in Precision Medicine........................................................................................................189
7.1 Chapter 7 Summary..................................................................................................................189
7.2 Introduction .............................................................................................................................189
7.3 Methods ....................................................................................................................................191
7.4 Results .......................................................................................................................................208
7.5 Discussion .................................................................................................................................216
7.6 Conclusion .................................................................................................................................218
7.7 Summary of Chapter 7..............................................................................................................219
Chapter 8 .........................................................................................................................................220
Discussion and Conclusions ...........................................................................................................220
8.1 Chapter 8 Summary ....................................................................................................................220
8.2 Contribution of this Thesis to Knowledge ..................................................................................220
8.3 Summary of Key Findings .........................................................................................................221
8.4 Summary of the Methodological Contribution of this Thesis ....................................................225
8.5 Limitations ..................................................................................................................................227
8.6 Implications of this Thesis ..........................................................................................................230
8.7 Conclusion ..................................................................................................................................232
9 References .................................................................................................................................234
Appendices ....................................................................................................................................262
Appendix 2.2: Search strategy .........................................................................................................283
Appendix 2.3: Summary of Included Systematic Reviews of Economic Evaluations of Precision Medicine ................................................................................................................................285
Appendix 2.4: References of all Included Economic Evaluations of Precision Medicine ..........295
Appendix 2.5: Location of Economic Evaluations of Precision Medicines (2007-2017) ........312
Appendix 2.7: Summary of Studies Which Discuss Capacity Constraints ................................315
Appendix 4.2: Participant Information Sheet ..................................................................................338
Appendix 4.3: Consent Form ...........................................................................................................341
Appendix 4.4: Interview Schedule for Service Providers ..............................................................342
Appendix 4.5: Interview Schedule for Service Commissioners ..................................................345

Word Count: 65,082
List of Tables

Table 2.1: Inclusion Criteria for Systematic Reviews of Economic Evaluations of Precision Medicine ........................................41
Table 2.2: Inclusion Criteria for Primary Economic Evaluations of Precision Medicine ..................................................43
Table 2.3: Capacity Related Terms Used in Keyword Search .................................................................44
Table 2.4: Summary of included studies using CHEERs checklist .............................................................54
Table 3.1: Incremental cost-effectiveness ratios of Mammaprint at different implementation levels ..................................................78
Table 3.2: VOImp (static) accounting for the baseline and final level of implementation of Mammaprint in a population of 10,000 women ..................................................................................81
Table 3.3: Current Value of Implementation (dynamic) for Mammaprint assuming constant or variable net monetary benefit ........................................................................................................83
Table 3.4: Actual Value of Implementation (dynamic) for Mammaprint assuming constant or variable net monetary benefit which improves uptake by 3 percentage points .............................................85
Table 4.1: A Typology of Barriers to the Introduction of Examples of Precision Medicine for NSCLC ..........................................................................................................................105
Table 5.1: Model input parameters for decision tree .........................................................................................137
Table 5.2: Model input parameters for the state transition Markov model .....................................................138
Table 5.3: Comparison of the deterministic results from this case study with published manufacturer estimates .........................................................................................................................................155
Table 6.1: Parameter Values for the Identified Capacity Constraints ..................................................................................170
Table 6.2: Value for each capacity constraint in each scenario ..............................................................................177
Table 6.3: The impact of the inclusion of capacity constraints on ICERs and NMB .................................................................................................................................179
Table 7.1: Cost and effect of the capacity-investment to Improve Awareness of ALK test commissioning ..................................................................................................................................................194
Table 7.2: Unit staff costs ......................................................................................................................................................198
Table 7.3: The Cost of Filling Vacancies in Pathology Laboratories with Different Levels of Centralisation ..................................................................................................................................................198
Table 7.4: Predicted Test Population Size in Future Years .........................................................................................200
Table 7.5: Parameter values for capacity-investment ‘Outreach Programme’ to increase commissioning awareness ........................................................................................................................................204
Table 7.6: Parameter values for a capacity-investment to centralise ALK testing .................................................205
Table 7.8: Parameter values for three combined capacity-investments in the provision of ALK testing .....................................................................................................................................................207
List of Figures

Figure 1.1: Example Incremental Cost-Effectiveness Plane .................................................. 29
Figure 2.1: Identification of Systematic Reviews of Economic Evaluations of Precision Medicine ...................................................................................................................... 46
Figure 2.2: Identification of Primary Economic Evaluations of Test and Treatment Precision Medicine .......................................................................................................................... 47
Figure 3.1: Current VOImp (dynamic) for Mammaprint assuming constant or variable net monetary benefit ......................................................... 82
Figure 3.2: Actual value of an implementation strategy which improves uptake by 3 percentage points ................................................................................................................................. 84
Figure 5.1: Decision tree¹ ........................................................................................................ 135
Figure 5.2: State-transition Markov model ............................................................................ 136
Figure 5.3: Weibull survival curve for progression free survival with crizotinib .................. 143
Figure 5.4: Log-normal survival curve for progression free survival with docetaxel ........... 144
Figure 5.5: Weibull survival curve for overall survival with crizotinib ............................... 145
Figure 5.6: Total incremental monetary net benefit of ALK testing to guide treatment with crizotinib or docetaxel by threshold level ................................................................. 152
Figure 5.7: Distribution of Estimated Incremental Costs and Benefits ................................. 153
Figure 5.8: Distribution of Annual Total Expected Net Monetary Benefit .......................... 154
Figure 6.1: Decision Tree to Incorporate Capacity Constraints ............................................ 165
Figure 6.2 Incremental Cost-Effectiveness Plane for the Results of the Model in the Presence and Absence of Capacity Constraints ................................................................. 181
Figure 6.3 Kernel Density Plot of Estimated Total NMB in the Presence and Absence of Capacity Constraints ................................................................. 182
Figure 7.1: Illustrative diffusion curves¹ .............................................................................. 201
Figure 7.2: Natural diffusion of EGFR and ALK testing .................................................... 202
Figure 7.3: The incremental monetary net benefit of a capacity-investment to improve commissioning awareness ......................................................... 209
Figure 7.4: The incremental net monetary benefit of a capacity-investment to promote centralising ALK testing ........................................................................................................ 210
Figure 7.5: The incremental net monetary benefit of a capacity-investment to improve pathology capacity .............................................................................................................. 211
Figure 7.6: The incremental net monetary benefit of three capacity-investments to improve the provision of ALK testing .................................................................................. 212
Figure 7.7: The incremental net monetary benefit of investing in commissioning awareness when ALK testing and crizotinib is not cost-effective ........................................... 213
Figure 7.8: The incremental net monetary benefit of investing in centralising ALK testing when ALK testing and crizotinib is not cost-effective ........................................... 214
Figure 7.9: The incremental net monetary benefit of investing in pathology capacity when ALK testing and crizotinib is not cost-effective ........................................... 215
Figure 7.10: The incremental net monetary benefit of three capacity investments to improve the provision of ALK testing when ALK testing and crizotinib is not cost-effective ........... 216
List of Appendices


Appendix 2.2: Search strategy

Appendix 2.3: Summary of Included Systematic Reviews of Economic Evaluations of Precision Medicine

Appendix 2.4: References of all Included Economic Evaluations of Precision Medicine

Appendix 2.5: Location of Economic Evaluations of Precision Medicines (2007-2017)

Appendix 2.6: Focus of Economic Evaluations of Precision Medicine (2007-2017)

Appendix 2.7: Summary of Studies Which Discuss Capacity Constraints


Appendix 4.2: Participant Information Sheet

Appendix 4.3: Consent Form

Appendix 4.4: Interview Schedule for Service Providers

Appendix 4.5: Interview Schedule for Service Commissioners
List of Abbreviations

ALK – Anaplastic lymphoma kinase
BRCA – Breast cancer susceptibility protein (type 1 or 2)
CADTH – Canadian Agency for Drugs and Technologies in Health
CBA – Cost-benefit analysis
CEA – Cost-effectiveness analysis
CUA – Cost-utility analysis
DAP – Diagnostics assessment programme
DES – Discrete event simulation
EGFR – Epidermal growth factor receptor
ERG – Evidence review group
EQ-5D – Euroquol 5 dimension
FISH – Fluorescent in situ hybridisation
HRQoL – Health-related quality of life
HTA – Health technology appraisal
HR – Hazard ratio
ICER – Incremental cost-effectiveness ratio
IHC – Immunohistochemistry
INMB – Incremental net monetary benefit
NHS – National Health Service
NGS – Next generation sequencing
NICE – National Institute for health and Care Excellence
NHB – Net health benefit
NMB – Net monetary benefit
NSCLC – Non-small cell lung cancer
PD-L1 – Programmed death ligand 1
PFS – Progression free survival
PRISMA – Preferred reporting items for systematic reviews and meta-analysis
QALY – Quality-adjusted life year
OS – Overall survival
RCT – Randomised control trial
ROS1 - Proto-oncogene tyrosine-protein kinase ROS
SF-36 – 36-item short form survey
TA – Technology appraisal
VOI – value of information
VOImp – value of implementation
WTP – Willingness to pay
## Glossary

<table>
<thead>
<tr>
<th>Term</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Access-limiting capacity</td>
<td>A capacity constraint which limits the number of patients who receive a health care intervention</td>
</tr>
<tr>
<td>Capacity constraint</td>
<td>Any factor which impedes or limits the amount of health status produced for a population of patients receiving specified interventions, or policies, provided by the healthcare system</td>
</tr>
<tr>
<td>Capacity investment</td>
<td>An implementation strategy targeted at reducing the impact of a capacity constraint on the provision of a health care intervention</td>
</tr>
<tr>
<td>Cost-benefit analysis</td>
<td>A comparison of the costs and benefits of alternative courses of action with the benefits measured using monetary units</td>
</tr>
<tr>
<td>Cost-effectiveness analysis</td>
<td>A comparison of the costs and benefits of alternative courses of action with the benefits measured in non-monetary units</td>
</tr>
<tr>
<td>Decision-analytic model</td>
<td>A mathematical approach to estimating the expected costs and benefits of alternative courses of action</td>
</tr>
<tr>
<td>Decision tree</td>
<td>A type of decision-analytic model which outlines the probabilities of a series of events occurring and the costs and outcomes of those events.</td>
</tr>
<tr>
<td>Diffusion</td>
<td>The natural rate at which a new healthcare intervention is adopted into the healthcare system in the absence of implementation strategies</td>
</tr>
<tr>
<td>Economic evaluation</td>
<td>A framework for the comparative analysis of alternative courses of action in terms of both their costs and consequences</td>
</tr>
<tr>
<td>Term</td>
<td>Definition</td>
</tr>
<tr>
<td>------------------------------------------------</td>
<td>-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Framework analysis</td>
<td>An approach to qualitative data analysis used in social science research characterised by the use of a matrix of themes and patients to help identify heterogeneity</td>
</tr>
<tr>
<td>Implementation strategy</td>
<td>An intervention which seeks to improve the provision, whether in terms of patient access or the quality of a healthcare intervention in the healthcare system</td>
</tr>
<tr>
<td>Incremental-cost effectiveness ratio</td>
<td>The ratio of the incremental costs and benefits produced by the potential introduction of a healthcare intervention</td>
</tr>
<tr>
<td>Markov model</td>
<td>A type of decision-analytic model involving cohorts of patients moving through a number of health states, with given probabilities and costs and benefits, at set time cycles</td>
</tr>
<tr>
<td>Mutation</td>
<td>An alteration in the sequence of deoxyribonucleic acid (DNA) which may result in the growth and development of cancer</td>
</tr>
<tr>
<td>Net health or monetary benefit</td>
<td>A measure of the cost-effectiveness of an intervention based on the improvement in total health or value provided by introducing the intervention</td>
</tr>
<tr>
<td>Precision medicine</td>
<td>An approach to disease management which uses information about the patient obtained through tests to target care and improve outcomes</td>
</tr>
<tr>
<td>Probabilistic sensitivity analysis</td>
<td>An approach for accounting for uncertainty in economic evaluations wherein each parameter is assigned a probability distribution and the model is re-run over many iterations with new parameter values drawn for each iteration</td>
</tr>
<tr>
<td>Quality-adjusted life year</td>
<td>A generic measure of health benefit incorporating health-related quality of life and length of life</td>
</tr>
<tr>
<td>Term</td>
<td>Description</td>
</tr>
<tr>
<td>-------------------------------------------</td>
<td>-----------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Quality-limiting capacity constraint</td>
<td>A capacity constraint which affects the costs or benefits provided by an intervention but does not reduce the number of patients who receive the intervention.</td>
</tr>
<tr>
<td>Thematic analysis</td>
<td>An approach to analysing qualitative data focussed on identifying patterns in the data.</td>
</tr>
<tr>
<td>Value of implementation</td>
<td>A framework to evaluate the cost-effectiveness of implementation strategies.</td>
</tr>
</tbody>
</table>
Abstract

Background

Precision medicine involves the use of tests, algorithms, or other information to target treatments to improve benefits or avoid harm for patients. Examples of precision medicine represent complex interventions and there are many barriers, or health system capacity constraints (hereafter ‘capacity constraints’), to their introduction into clinical practice. The assumptions underlying economic evaluations of healthcare interventions do not currently allow for the incorporation and measurement of the impact of capacity constraints. The overall aim of this PhD was to identify and quantify the impact of including capacity constraints in decision-analytic model-based cost-effectiveness analysis to better inform resource allocation decisions and the introduction of precision medicine into clinical practice.

Methods

This thesis comprised six empirical chapters using different methods. A systematic review of systematic reviews (meta-review) was used to identify existing economic evaluations of precision medicine to determine if and how these studies quantified the impact of capacity constraints. Static and dynamic value of implementation methods were adapted to allow for the varying marginal costs and benefits which may arise due to health system capacity constraints and these methods were applied to a case study in breast cancer. To develop a typology of barriers to the introduction of examples of precision medicine, qualitative interviews were conducted with stakeholders in the implementation of examples of precision medicine in non-small cell lung cancer. A case study example of precision medicine was selected based on these interviews and a base case decision analytic model, replicating a published technology appraisal conducted by the National Institute for Health and Care Excellence (NICE), was created to evaluate the cost-effectiveness of immunohistochemistry (IHC) and fluorescent in-situ hybridisation (FISH) testing for anaplastic lymphoma kinase (ALK) alterations to guide treatment with crizotinib or docetaxel. Three capacity constraints were selected based on the qualitative interviews and these were incorporated into the base case model. The impact of these capacity constraints was quantified using static value of implementation methods. Hypothetical investments to reduce the impact of the capacity constraints were proposed and their potential value was determined using dynamic value of implementation methods.
Results

The meta-review identified 45 previous reviews of economic evaluations of precision medicine from which a sample of 222 economic evaluations focusing on ‘test-treat’ interventions were selected. Of these, 33 studies qualitatively discussed potential capacity constraints and nine studies attempted to quantify the impact of capacity constraints. It was identified that capacity constraints may impact the marginal costs and benefits of an intervention and therefore its cost-effectiveness. When the static and dynamic value of implementation methods were adapted to allow for this varying marginal cost-effectiveness, it was identified that the way in which examples of precision medicine are implemented would be critical to ensuring a net benefit was achieved for society.

From the qualitative interviews, a typology of 17 barriers to the introduction of examples of precision medicine in NSCLC was created. A base case economic evaluation of one of these examples of precision medicine, ALK testing, provided an estimate of the cost-effectiveness of the example in the absence of capacity constraints. The estimated ICER of £38,468 was similar to that produced in the original health technology assessment of crizotinib (£41,554). In the absence of capacity constraints, ALK testing and treatment with crizotinib offered a potential total net monetary benefit of £6,373,887 per year. Incorporating three capacity constraints into this model reduced the estimated cost-effectiveness and total net monetary benefit of ALK testing to £41,413 per QALY gained and £1,084,473 respectively. Investing in a strategy to overcome these constraints was found to have a significant incremental net monetary benefit to society (£6,247,486 over six years), particularly in scenarios when the intervention was not cost-effective at the time of NICE technology appraisal (£36,431,395 over six years).

Conclusion

Health system capacity constraints can have a significant impact on the cost-effectiveness and net monetary benefit of examples of precision medicine. These impacts and the value of investing to reduce the impact of capacity constraint can be quantified using static and dynamic value of implementation methods which allow for varying marginal costs and benefits.
Declaration

This dissertation is entirely the result of my own research. No portion of the work referred to in the thesis has been submitted in support of an application for another degree or qualification of this or any other university or other institute of learning.

Chapters 1, 5, 6, 7 and 8 were written as traditional thesis chapters. The first drafts were written by me and re-written following feedback from my supervisors.

Chapters 2, 3, and the methods section of chapter 4 report published peer-reviewed journal articles.


I developed the idea for the review and the search strategy. I screened abstracts, extracted data, wrote the narrative analysis, and wrote the first draft of the manuscript. All authors contributed towards the further drafting and editing of the manuscript.


All authors were involved with developing the idea for the paper. I adapted the equations to develop the method and applied it to the existing case study. I wrote the first draft of the manuscript. All authors contributed towards the further drafting and editing of the manuscript.


This paper presents the methods section of Chapter 4. I developed the idea for the paper. I created the interview schedule, conducted the interviews, analysed the data, and wrote the first draft of the manuscript. All authors contributed towards the further drafting and editing of the manuscript.
Copyright Statement

I. The author of this thesis (including any appendices and/or schedules to this thesis) owns certain copyright or related rights in it (the “Copyright”) and s/he has given The University of Manchester certain rights to use such Copyright, including for administrative purposes.

II. Copies of this thesis, either in full or in extracts and whether in hard or electronic copy, may be made only in accordance with the Copyright, Designs and Patents Act 1988 (as amended) and regulations issued under it or, where appropriate, in accordance with licensing agreements which the University has from time to time. This page must form part of any such copies made.

III. The ownership of certain Copyright, patents, designs, trademarks and other intellectual property (the “Intellectual Property”) and any reproductions of copyright works in the thesis, for example graphs and tables (“Reproductions”), which may be described in this thesis, may not be owned by the author and may be owned by third parties. Such Intellectual Property and Reproductions cannot and must not be made available for use without the prior written permission of the owner(s) of the relevant Intellectual Property and/or Reproductions.

IV. Further information on the conditions under which disclosure, publication and commercialisation of this thesis, the Copyright and any Intellectual Property and/or Reproductions described in it may take place is available in the University IP Policy (see http://documents.manchester.ac.uk/Doculnfo.aspx?DocID=24420), in any relevant Thesis restriction declarations deposited in the University Library, The University Library’s regulations (see http://www.library.manchester.ac.uk/about/regulations/) and in The University’s policy on Presentation of Theses.
Rationale for Journal Format Submission

This thesis is written using the journal format in which published papers comprise some of the chapters. These published papers report research conducted as part of the PhD programme and so fit within the structure of the thesis. A chapter summary introduces each chapter and explains how the chapter fits into the overall structure of the thesis.
Publications

This PhD comprises eight chapters. To date, elements of the content presented in chapters 2, 3, and 4 have been published.


I developed the idea for the review and the search strategy. I screened abstracts, extracted data, wrote the narrative analysis, and wrote the first draft of the manuscript. All authors contributed towards the further drafting and editing of the manuscript.


All authors were involved with developing the idea for the paper. I adapted the equations to develop the method and applied it to the existing case study. I wrote the first draft of the manuscript. All authors contributed towards the further drafting and editing of the manuscript.


This paper presents the methods section of Chapter 4. I developed the idea for the paper. I created the interview schedule, conducted the interviews, analysed the data, and wrote the first draft of the manuscript. All authors contributed towards the further drafting and editing of the manuscript.

Two additional papers will be prepared for submission to journals for publication:


Working title: Quantifying the Impact of Capacity Constraints Using Value of Implementation Methods. Possible Target Journal: Medical Decision Making. This will report the study described in Chapter 6.


Working title: Valuing Investments in Health System Capacity to Support the Implementation of an Example of Precision Medicine. Possible Target Journal: Medical Decision Making. This will report the study described in Chapter 7.
Other papers on topics related to this PhD published during the completion of this thesis (between 2016 and 2020):


Other papers on topics not directly related to this PhD published during the completion of this thesis (between 2016 and 2020):


Wright SJ, Ulph F, Lavender T, Dharni N, Payne K. Understanding Midwives’ Preferences for Providing Information About Newborn Bloodspot Screening. Medical Decision Making Policy & Practice. 2017;3(1)


Acknowledgements

I would like to thank my supervisors, Professor Katherine Payne, Professor William Newman and Dr Gavin Daker-White for their support throughout this entire PhD programme, from helping to develop the initial proposal for the studentship through to the writing of this final thesis. Their guidance has improved this work immeasurably. I would also like to thank Dr Jon Gibson for acting as my advisor during this PhD. Thanks also go to Professor Fiona Blackhall and Dr Yvonne Summers for their advice on the clinical aspects of treating non-small cell lung cancer and to Dr Nathan Proudlove for acting as a sounding board for some of the initial ideas for this body of work. In addition I would like to express my gratitude to the Wellcome Trust for funding this PhD project.

A special thanks to everyone at the Manchester Centre for Health Economics for their support, advice, and feedback on the work contained in this thesis over the last four years. Thanks also to those who have provided feedback on work from this thesis which has been presented at national and international conferences including the Society for Medical Decision Making and International Society for PharmacoEconomics and Outcomes Research.

Finally, I would like to thank my friends and family for their continued support. In particular I would like to thank my wife Sophie and son Jamie for providing me with love and keeping me motivated throughout my studies.
The Author

The author, Stuart Wright, was award a Batchelor of Science in Economics degree by the University of Kent in 2010 and a Master of Science in Health Economics degree by the University of East Anglia in 2013.

Stuart began working as a research associate in health economics at The University of Manchester in October 2013 and was awarded a Wellcome Trusts Society & Ethics Doctoral Studentship in 2016 to complete this PhD. In 2019 he began a role as a research fellow in heath economics at The University of Manchester.
Chapter 1

Introduction

1.1 Thesis Summary

The overall aim of this PhD was to identify and quantify the impact of including health system capacity constraints (hereafter ‘capacity constraints’) in decision-analytic model-based cost-effectiveness analyses to better inform resource allocation decisions and the introduction of precision medicine into clinical practice.

This thesis is underpinned by extra-welfarist economic theory and uses methods to quantify the impact of including capacity constraints in decision-analytic model-based cost-effectiveness analysis. In this thesis, capacity constraints are defined as any factor which limits the maximum amount of health produced for patients by the health system, by impeding the potential health producing actions of the system (see also section 1.8). The focus of this thesis is the potential impact of such capacity constraints on conducting economic evaluation of examples of precision medicine. There is no unified definition of precision (also sometimes called stratified or personalised) medicine but in practice to date it refers to using some test-treat combination to target an intervention (see also section 1.2). In this thesis, it is assumed that precision (stratified or personalised) medicine is underpinned by the premise that it is feasible to identify known heterogeneity using a specific test or algorithm in patient populations to guide patient care to improve health and wellbeing.

This thesis used a combination of different methods involving systematic literature reviews, qualitative, and quantitative methods. The thesis comprises eight chapters, of which, six are empirical chapters. Chapter 1 provides an outline of this thesis by introducing the key concepts of precision medicine, the role of economic evaluations, and capacity constraints. The aim, objectives and research question are outlined alongside the study design and method which will be used to address these in each of the six empirical chapters (Chapters 2 to 7).

1.2 Precision Medicine

Precision medicine, also commonly referred to as personalised or stratified medicine, is “an emerging approach for disease treatment and prevention that takes into account individual variability in genes, environment, and lifestyle for each person” (1). Implementing precision medicine involves identifying information about a patient and/or their disease, such as “prognostic markers, predictors of toxicities and any parameter such as environmental and lifestyle factors”
which predict potential improved or reduced response to treatments or higher risk of side effects (2). A strategy is then developed using this information, via a test or algorithm, to target treatments to improve outcomes or wellbeing for patients.

Treatment for patients can be targeted based on tests identifying a wide range of biomarkers including genes, metabolites and proteins (3). This thesis focuses on genomic tests which seek to identify changes in the genes of a patient, or in the case of cancer the tumour, in order to guide treatment decisions. Genomic alterations which may impact treatment decision for patients can be divided into two groups: germline or somatic (4). Germline genetic alterations are passed on from parents to their children whereas somatic mutations occur in the body after conception and are not passed on to future generations (5,6). For example, testing patients receiving warfarin for germline cytochrome P450 2C9 (CYP2C9) gene variations can guide dosing with the aim of reducing major bleeding complications (7). In an alternative example, colorectal cancer patients who have tumours with somatic KRAS gene mutations experience poorer outcomes when treated with the drug cetuximab so testing for such mutations can be useful in guiding treatment (8).

Targeting medicines, using test-treat strategies, to those in a population more likely to benefit or less likely to be harmed, may in theory offer clinical and economic benefits. These potential clinical and economic benefits could be realised by: targeting treatments to responders; targeting treatments to avoid adverse-events; health care cost-savings from avoiding treating patients who are unlikely to benefit from a medicine.

One example of a precision medicine where testing is used to identify patients who are more likely to respond to a treatment is EGFR mutation testing and gefitinib. Gefitinib was originally approved by the United States Food and Drug Association (FDA) in 2003 for all patients with non-small cell lung cancer who had not responded to standard chemotherapy (9). However, only 14% of patients enrolled in early trials responded to the treatment and two subsequent larger trials showed no benefit in response rate, time to progression, or overall survival compared to chemotherapy (10). As such, gefitinib was later withdrawn from the market (9). In 2005 it was found that response to gefitinib was correlated to epidermal growth factor receptor (EGFR) gene mutations in non-small cell lung cancer (NSCLC) tumours (11). The ability to target gefitinib to those who would experience benefit led to its approval in the European Union from 2009 (12). The ability to restrict the use of expensive targeted treatments to those patients in a population who will benefit from them, while providing an alternative appropriate treatment to those patients identified as non-responders, has been highlighted as one of the key economic benefits of precision medicine (13).
An alternative benefit presented by some examples of precision medicine is the use of testing to predict which patients may experience adverse events, and therefore reduced health outcomes, from taking a medicine. For example, genotyping can be used to identify patients with low or no thiopurine methyltransferase (TPMT) enzyme activity (14,15). Such patients are more likely to experience the adverse reaction of neutropenia when they are prescribed azathiopurine for conditions such as rheumatoid arthritis and by identifying them before treatment begins, an alternative therapy can be offered. Other examples of precision medicine which are used to avoid adverse events include HBA B*5701 testing before treatment with abacavir for patients with human immunodeficiency virus (HIV) or UGT1A1 testing before treatment with the chemotherapy agent irinotecan (16,17).

One perceived economic benefit of precision medicine has been the potential for cost savings from improved diagnosis and management of diseases using prognostic biomarkers. Multi-gene assays such as Mammaprint (18) and Oncotype DX (19) can help to predict the risk that currently remising breast cancer will recur, enabling lower risk patients to be treated only with endocrine therapy rather than costlier and more harmful chemotherapy (20,21). Furthermore, cost savings may also be realised by preventing the cancer from advancing and becoming harder to treat.

In reality, the experience of introducing many examples of precision medicine into clinical practice has been one of increasing costs to healthcare systems. Many targeted therapies, particularly those in cancer, have a high monthly cost and may be significantly more expensive than chemotherapy or immunotherapy over a number of years. For example, the drugs afatinib, ceritinib, and osimertinib for treating patients with EGFR, ALK, and EGFR T790M mutation positive non-small cell lung cancer respectively have a monthly cost of £2,023, £2,757, and £5,770 (22–24). Alternatively, in applications such as cardiovascular disease, the large volume of people who could benefit from pharmacogenetic treatment, such as genotype-guided warfarin dosing, may mean that there is a significant budget impact even when the cost of a single test is small (25).

### 1.3 Opportunity Cost

A key challenge faced by healthcare decision makers at national, regional and local levels is how finite budgets and resources should best be allocated to maximise outcomes for patients. When funding and resources are provided for one intervention, they are not available for use in other areas which could also provide benefit for patients. The potential benefit foregone by the use of resources for one purpose as opposed to others is known as the opportunity cost (26). Opportunity cost is the key concept underpinning the economic evaluation of healthcare interventions.
It is important that decision makers, representing different levels within a healthcare system (national, regional, hospital level) (27), weigh the opportunity costs of their decisions in order to determine whether greater benefits could have been achieved elsewhere. If it is assumed that the overall budget of a health system is fixed, then any funding of new interventions takes resources away from existing interventions. The introduction of new approaches to treatment, such as precision medicine, therefore potentially harms other patients in the health system by reducing funding for other interventions and the balance of benefits and harms involved in allocating and reallocating resources must be carefully evaluated (28).

1.4 The Role of Economic Evaluation

In order to compare the opportunity cost of funding decisions in the NHS, a comparison of the benefits as well as the costs of different health care interventions must be made. Opportunity cost can be identified, measured and valued using the methods of economic evaluation. Decision making regarding the types of services and interventions which are provided in a health system, as well as which additional interventions should be introduced, can therefore be assisted by information from economic evaluations of healthcare interventions. Economic evaluations present a framework for “the comparative analysis of alternative courses of action in terms of both their costs and consequences” (29).

Implementing economic evaluation poses problems as there are many definitions of healthcare benefit. There are two main normative frameworks for the identification, measurement and valuation of benefit in economic evaluation: welfarism and extra-welfarism. A welfarist approach to economic evaluation place value on interventions based on the utility they provide to patients. While utility is a broad term and has been used in many ways, Brouwer et al., suggest that the most commonly used interpretations are that utility is satisfaction of preferences or a desirable state of consciousness (30). When an intervention is introduced for a group of patients, they may experience increased utility from their improved health. However, funding may have to be withdrawn from interventions for other patients in the health system and this group would lose utility. In the welfarist approach healthcare interventions are valuable if after their introduction, the net, aggregated utility of patients in the health system has increased. The measurement of utility is most commonly achieved by asking patients the maximum amount they were willing to pay for a health intervention (WTP) (31). The evaluation of interventions based on their costs and monetary benefits is known as cost-benefit analysis (CBA) (32). CBA is used infrequently in practice due to technical and ethical challenges (33,34).
The alternative normative approach to identifying, measuring and valuing outcomes in economic analyses is the extra-welfarist approach. The ‘extra’ in extra-welfarism refers to use of an outcome other than utility, for example a clinical outcome or a generic measure of health benefit. For example, progression free survival is a common outcome measure in clinical trials of treatments for NSCLC (35–37). New interventions which treat cancer could be evaluated based on the amount they would cost per month of progression free survival they provide. This type of economic evaluation is known as a cost-effectiveness analysis (CEA) and is characterised by the comparison of costs and a relevant clinical outcome measure (38). A disadvantage of CEA is that clinical outcomes, such as progression free survival, are often specific to interventions in a given area. As such the cost-effectiveness of a cancer intervention which mainly extends progression free survival could not be compared to a treatment for cardiovascular disease which reduces the risk of stroke. A generic outcome measure which can be used in all areas of health is therefore required when the focus is on producing the best allocation of health resources across all disease areas. The most commonly used of such outcomes in health is the quality-adjusted life year (QALY). The QALY is a composite measure which combines changes in health-related quality of life (HRQoL), which have been preference weighted to form health utilities, and duration of life (39). As the concept of HRQoL is broad, the QALYs gained from interventions can be calculated and compared for health care interventions in a wide range of areas.

Typically HRQoL is captured using a multi-attribute questionnaire such as the EQ-5D or SF-36 (40,41). These ask questions about a patient’s current experience of a range of health domains including: pain, anxiety or depression, social functioning and mobility. Responses to these questions are then converted into a cardinal value known as a health utility bounded at a value of one for full health and zero for death (42). Values less than 0 are possible and represent health states which are worse than death. These health utilities are generated by determining the preferences of large samples of the public for different aspects of the underlying HRQoL measure. For example, in the UK, changes in the pain domain of the EQ-5D 3L carry the most weight in determining health utility while changes in the usual care domain carry the least weight (42).

The changes in a patient’s health utility are summed over the duration of the patient’s life to obtain the QALYs gained by the use of the intervention (43). Interventions can then be compared on the basis of the cost per QALY generated using an economic evaluation known as cost-utility analysis (CUA). CUA is the approach to economic evaluation which is recommended by NICE in the UK, although it is commonly confusingly referred to as cost-effectiveness analysis in the methods guides (44). The only difference between cost-effectiveness analysis and cost-utility analysis is the use of an outcome measure incorporating HRQoL. Within this thesis, the more general term of cost-effectiveness analysis (CEA) will be used to refer to the cost-utility approach.
1.5 Generating Economic Evidence

Two methods (or vehicles (45)) are available for collecting and comparing cost and outcome data for the purpose of cost-effectiveness analysis. Cost and QALY data may be collected as part of prospective trials, randomised or non-randomised, investigating the effectiveness of new interventions compared with current practice (46). Such data will benefit from the quality of data collection associated with individual-patient level prospective design. However, trials often only run for a set period of time whilst the costs and benefits of treatment may extend into the future (45,46). Furthermore, trials are often designed so that enough patients are recruited to show significant differences in clinical outcomes but not necessarily economic ones (47). It may therefore be difficult to identify whether there are significant improvements in QALYs in a single trial. Trials may also focus on the clinical benefit of an intervention in a specific patient population.

An alternative, or potentially a supplement, to using a trial to collect economic data is the use of decision-analytic models (48). A decision-analytic model is a “formal quantified comparison of health technologies, synthesising sources of evidence on costs and benefits, in order to identify the best option for decision makers to adopt” (48). A decision-analytic model aims to create a simplified representation of the experiences of patients who are treated with two or more different health interventions with the aim of comparing differences in expected costs and outcomes with the purpose of answering a specified decision-problem (49). The added flexibility of decision-analytic models can be used to extrapolate data beyond the endpoints of a prospective study, combine data from different trials or different data sources, and compare interventions with different comparators which were not used in the original trials (45,50). Decision-analytic models can also be used to explore the effects of different scenarios, for example changes in the costs of different interventions and their impact on cost-effectiveness (51).

Different types of decision-analytic models are available depending on the disease and intervention being evaluated as well as whether patients are modelled as a cohort or at the individual level (48). In cohort-based models, the eligible population of patients is modelled using aggregate level data. The simplest type of cohort-based decision-analytic model is the decision tree. In a decision tree, an initial decision is made between treating patients with different interventions and then the future consequences of treatment are mapped out in increasing ‘branches’ (of the tree) (50). Each potential consequential event occurs with a given probability and costs and QALYs are attached to the final event in each branch. The expected costs and QALYs of each treatment are compared to determine the cost-effectiveness of the intervention. As decision trees are simple and events occur
in a sequence, they function best in diseases which occur over short time frames and where treatment progresses in a linear manner (52).

An alternative type of cohort-based decision analytic model is the state transition Markov model. A state transition Markov model (hereafter Markov model) features a selection of health states, potentially including healthy and dead, which patients transfer between over a given time period (53). Patients may return to health states which they have previously left which make a Markov model more suitable for capturing the progression of chronic diseases where patients can repeatedly move between healthy and unhealthy states (54). As such, Markov models are used extensively in evaluating cancer interventions where patients can have the disease and go into remission or can be in remission and suffer a relapse of their disease (55–57). Patients transition through the Markov model over a series of time periods and suffer a QALY loss (health disutility) and costs of different sizes whilst they remain in disease related health states. To compare treatments, the same model structure is often used but the probability of moving between states, health disutilities and costs of being in given states are differed according to health data. The health disutilities are aggregated to form QALY estimates and compared across interventions to calculate the QALY gain from a new intervention (50). The incremental cost is calculated to gain an estimate of the incremental cost per QALY for the new intervention assuming the relevant perspective, such as that of the health system. A number of decision-analytic models evaluating the cost-effectiveness of precision medicine for NSCLC feature Markov models (58,59). In addition, other studies use linked decision trees and Markov models to evaluate the cost-effectiveness of mutation testing and precision medicine (60,61).

While cohort level models are useful for many disease areas, they also have limitations. For example, a patients’ experience of a disease or care pathway may depend on a variety of factors including their demographic information, history of disease or past treatment. Other more complex models, that represent a population of patients at the individual-level, are available including discrete event simulation. Markov models are limited by being “memoryless”; the previous experience of an individual in the pathway represented by the decision-analytic model cannot affect their future probability of moving between health states, health utility, or costs (50,62). For example, in a cancer-specific Markov model the probability that someone suffers a given side effect of treatment could not depend on whether they had suffered that side effect before or for a different medication. Probabilities, disutilities and costs are also average values for the cohort and cannot depend on individual characteristics (48). Furthermore, Markov models follow individuals through health states rather than events which can be conceptually awkward for some applications (63). Discrete event simulation removes some of these constraints by allowing individuals to be followed through the decision-analytic model as they experience different health related events.
However, this added flexibility comes at a cost of significantly higher computational requirements and a lower degree of transparency for decision makers.

1.6 The Cost-Effectiveness Threshold

After identifying, measuring, valuing and comparing the costs and QALYS of health care interventions, decision makers must apply rules (decision-rules) to determine which should be incorporated in the health care system. The incremental costs and QALYs estimated from an economic evaluation are commonly plotted on a graph called an incremental cost-effectiveness plane (figure 1.1).

**Figure 1.1: Example Incremental Cost-Effectiveness Plane**

If the point plotted from the incremental costs and benefits of the intervention lies in the South East corner of the plot, the intervention is said to dominate the comparator. This is because the intervention is both cheaper and more effective than the comparator and should therefore be adopted into the health system. Conversely, interventions in the North West quadrant of the plane are dominated by the comparator as they are both more expensive and less effective. The decision as to whether an intervention in the North East (more costly but more effective) or South West (less costly but less effective) should be adopted is less clear.

To aid in the decision as to whether to adopt interventions with higher costs and benefits to the comparator, a ratio of the benefits to costs is created. This is known as the incremental cost-
effectiveness ratio (ICER) and in CUA is presented as the cost per QALY gained. Lower ICER values are preferred for interventions in the North East quadrant of the cost-effectiveness plane while higher ICER values are preferred for interventions in the South West quadrant.

A decision rule is required to determine whether an intervention is cost-effective based on the estimated ICER. The threshold approach is the most broadly used decision-rule in evaluating the relative cost-effectiveness of new interventions. The threshold is a benchmark cost per QALY which is deemed acceptable. This ratio is compared against the threshold cost per QALY and if it is lower than the new intervention it is deemed to be cost-effective (66–68).

The threshold is most commonly interpreted as the marginal cost of producing a QALY in the health system. If a QALY can be produced at a lower cost by a new intervention than from the least efficient existing intervention in the health system, then it should be introduced (68). This is because the total number of QALYs produced by the health system can be increased by this substitution. As a result of the introduction of the more cost-effective intervention, the threshold should then fall as on average the health system has become more efficient at producing QALYs and so it will be harder to make further gains in the future (69).

The National Institute for Health and Care Excellence (NICE) works under the remit of NHS England to assess whether specified new interventions have sufficient evidence of clinical and cost-effectiveness to be adopted into clinical practice. It has three programmes relevant to the evaluation of precision medicine: the NICE diagnostics assessment programme (NICE DAP) which includes evaluations of mutation testing, the highly specialised technologies programme (NICE HST) which considers treatments for very rare conditions, and the NICE technology appraisal programme (NICE TA) which cover new medicines (44,70). NICE uses a stated threshold of £20,000 to £30,000 per QALY to evaluate new health technologies (44). For some treatments, deemed to be offering end-of-life options for specified diseases such as NSCLC, with small patient groups who have a life expectancy of less than 24 months and where a treatment may extend life by more than 3 months, a higher threshold may be used (71). A higher threshold again may be used for medicines for patients with very rare diseases under the highly specialised technologies programme (72).

1.7 Assumptions of Decision-Analytic Model-Based Cost-Effectiveness Analysis

A number of key assumptions underpin using the results of a decision-analytic model-based cost-effectiveness analysis with a decision-rule applying the threshold approach to informing health care resource allocation decisions. These assumptions include that the healthcare interventions
being evaluated are completely divisible and exhibit constant marginal costs and QALYs (66,73,74). The divisibility of interventions implies that they can be funded completely or in part such that only a proportion of the total patient population receives them (64). The assumption of constant marginal costs and QALYs, otherwise known as ‘constant returns to scale’, requires that these marginal effects of treating additional patients are always the same regardless of the number of patients treated (75).

There are a number of reasons why the assumption of constant returns to scale may not hold in practice. New interventions may require a significant initial investment in capital which means that it is not cost-effective to treat a limited number of patients but that when these costs are shared over a greater patient population the intervention does constitute an efficient use of resources (75,76). There may be (dis)economies of scale which reduce or increase the marginal cost per patient as implementation increases (77,78). For example, it may be that manufacturers offer discounts for purchasing larger quantities of reagents for a diagnostic test. Similarly, economies of scope may mean that resources can be shared over multiple interventions, thereby reducing the marginal cost for a specific individual intervention (77). Alternatively, stretching underfunded resources over multiple interventions may reduce the efficiency of individual programmes. There may be a learning curve for a new technology which means that clinicians’ experience and, therefore, patient health outcomes increase with increased use of the intervention (79). Prioritisation of patients who experience increased QALYs from an intervention may lead to decreasing marginal health benefits as the intervention is rolled out to the whole patient population (80).

The assumption of the divisibility of healthcare interventions may be violated by decision makers on an ethical basis. This violation may occur because it is deemed inequitable to provide a new and more effective intervention to some members of the population while withholding it from others (81). The avoidance of such situations is one of the key aims of NICE and the requirement that medicine recommended for use in the NHS should be made available for all patients within 3-months of the NICE decision effectively breaks the assumption of divisibility (82). However, treating interventions as indivisible means that the total size of the programme in terms of the resource required for its implementation becomes a factor in decision making. This is because the implementation of particularly large healthcare programmes in full may require disinvestment from a number of smaller interventions which each have different levels of cost-effectiveness (83). This situation may cause problems when the incremental cost-effectiveness of the new intervention is close to the threshold, in which redistribution of resources across the new and existing programmes may provide the optimal solution rather than simply adopting the new intervention as dictated by the threshold approach (81).
1.8 Health System Capacity Constraints

The role of decision-analytic model-based cost-effectiveness analysis in informing resource allocation decisions in health care is clear (84). However, current decision-analytic model-based economic evaluations do not consider constraints within the capacity of health systems. Brennan et al (2006) suggest the importance of capturing the impact of health system capacity constraints when describing a taxonomy of decision-analytic models: ‘Inaccurate cost-benefit assessments can result from ignoring the interactions between service capacity decisions, waits and health outcomes’ (48). There is no consensus definition of health system capacity constraints. A title search of the Embase and MEDLINE databases conducted in OVID for the term capacity constrain* resulted in 40 papers being identified, none of which included a definition of what constituted a capacity constraint.

In this thesis a definition of health system capacity constraints was constructed by combining the dictionary definitions of ‘capacity’ and ‘constraints’: (see also Chapter 2, section 2.3). The Oxford English Dictionary (OED) defines capacity as ‘the amount that something can produce’ (85). This definition implies that it is necessary to be clear what is being produced. The capacity of a healthcare system could therefore be defined as its ability to produce some defined output. In keeping with the extra-welfarist normative underpinning, assumed by decision-making bodies such as NICE, the relevant output of a healthcare system has been defined as ‘health status’ measured using the QALY. The definition of a constraint is “something that controls what you do by keeping you within particular limits” (86). Combining the definitions of capacity and constraint can be used to propose a working definition of a capacity constraint in a healthcare system which has the goal of maximising health status:

“Any factor which impedes or limits the amount of health status produced for a population of patients receiving specified interventions, or policies, provided by the healthcare system.”

1.9 The Importance of Understanding the Impact of Constraints in Health System Capacity to the Provision of Precision Medicine

Precision medicine, as defined for the focus of this thesis, relies on the combination of two components: (i) a test, which identifies relevant heterogeneity between patients (ii) disease management strategies (‘treatments’) for the different groups of patients identified. Precision medicine can therefore be conceptualised as requiring a test-treat intervention. All examples of precision medicine are therefore complex interventions which, in addition to having a number of
interacting components, also pose issues in the “number and difficulty of behaviours required by those delivering or receiving the intervention”, the number of groups or organisations involved in providing the intervention, and the degree of tailoring of the intervention allowed (87).

The issue of the capacity of health systems to move precision medicine from the laboratory to clinical practice has been identified across disease areas. A 2013 report into the commissioning of molecular diagnostics identified that existing laboratory services may require additional resources and investment to satisfy demand for testing for precision medicines (88). Other key barriers to providing testing capacity included a lack of clarity regarding how local testing was commissioned, a need for quality assurance and maintaining the quality of NHS testing which was not regulated.

Two published qualitative studies have sought to investigate which barriers may interfere with the introduction of precision medicine. These two studies identified a key factor was a lack of clinical guidelines or clinician knowledge regarding the proper use of precision medicine and testing may limit patients’ access to such treatments (89,90). A number of constraints were rooted in a lack of healthcare resources and these included: long test turnaround times (89,90); additional time pressures for health professionals (91); and a need for geneticist support after testing (91). The additional cost of testing (90–92) and treatment (90,92) were also key barriers to implementing precision medicine.

1.10 Potential Implications of Capacity Constraints for the Economic Evaluation of Precision Medicine

In theory, using test-treat strategies to implement precision medicine represent healthcare interventions which can potentially produce additional health compared with current practice. Decision-analytic model-based cost-effectiveness analysis may be used to provide the required evidence to determine whether the introduction of a new approach to deliver precision medicine is likely to be cost-effective. Under the paradigm of an extra-welfarist approach a cost-effective intervention is one that produces a net gain in the total societal health produced by the health system.

Capacity constraints in the health system may impede the provision of a new precision medicine to all potentially eligible patients. The presence of such constraints could potentially contradict NICEs requirement that all patients have access to a medicine within 3-months of approval leading to concerns over equity of access to treatments. Patients not receiving the precision medicine may miss out on potential improvements in their quality and length of life when compared with those patients able to receive the intervention. Taking an economic perspective means that if some
patients do not receive a potentially cost-effective intervention then the maximum total incremental societal health gain from introducing the intervention may not be achieved (93). In other words, if more patients received the intervention then the total number of quality-adjusted life years produced by society would be higher.

Capacity constraints may also have a significant impact on the cost-effectiveness of an intervention if they impact on the cost or outcomes produced by the intervention. This situation may occur if the constraint increases the cost of providing the intervention to patients or reduces the expected benefit for each patient (94). For example, if clinicians need to learn new techniques to effectively use a precision medicine but a capacity constraint means that they see few patients each year, the benefits per patient and therefore the cost-effectiveness of the intervention may be reduced. In such cases capacity constraints will impact the size of the change in societal health but also mean that there could be a net reduction in the total health produced by the system instead of a net gain.

Given that health system capacity constraints may limit patient access to potential beneficial examples of precision medicine, and that they may also impact the cost-effectiveness of such interventions, evidence is needed as to the degree to which current economic evaluations of precision medicine take these constraints into account in their analyses. If current decision-analytic model-based economic evaluations do not account for health system capacity constraints, then the potential consequences of this omission should be evaluated using appropriate methods.

The overall aim of this PhD was to identify and quantify the impact of including health system capacity constraints (hereafter ‘capacity constraints’) in decision-analytic model-based cost-effectiveness analysis to better inform resource allocation decisions and introduction of precision medicine into clinical practice.

1.11 Research Questions

The aim of this PhD was specified by five research questions:

- Have published decision-analytic model-based cost-effectiveness analyses of examples of precision medicine accounted for health system capacity constraints?
- How could health system capacity constraints be included in decision-analytic model-based cost-effectiveness analysis of an example of precision medicine?
- What are the health system capacity constraints which may impede the introduction of examples of precision medicine into clinical practice?
• How can the impact of health system capacity constraints be measured in decision-analytic model-based cost-effectiveness analyses of an example of precision medicine?
• Can an economic perspective be used to evaluate the impact of the removal of health system capacity constraints to the introduction of an example of precision medicine?

1.13 Thesis Outline

This PhD takes an economic perspective underpinned by an extra-welfarist view. The empirical studies in this PhD used a both qualitative and quantitative methods to address the five stated research questions (section 1.11) and associated aims for each empirical study. There are six empirical chapters (Chapters 2 to 7).

Chapter 2 aimed to identify a sample of published economic evaluations of precision medicine and describe if, and how, these economic evaluations had qualitatively discussed and quantitatively accounted for capacity constraints in the analysis. The chapter presents a systematic review of systematic reviews (a meta-review) of economic evaluations of precision medicine. A sample of existing economic evaluations of precision medicines were identified and a key word search of terms relevant to health system capacity was conducted to identify studies which had discussed potential capacity constraints and those that had attempted to quantitatively account for constraints in the analysis. The result of the study were used to identify the optimal approach to including capacity constraints in economic evaluations of precision medicine.

Chapter 3 aimed to describe the application of value of implementation methods, using the static and dynamic forms to illustrate the potential impact of accounting for varying marginal costs and benefits on the relative cost-effectiveness and value of an exemplar of precision medicine. The traditional value of implementation method is adapted to allow for varying marginal cost and benefits which may directly arise from capacity constraints or indirectly interact with them. The method is applied to a case study example of precision medicine in breast cancer.

Chapter 4 aimed to explore the type and extent of barriers experienced by service providers and service commissioners when introducing precision medicine for NSCLC for relevant individuals within specific patient populations. The chapter outlines a qualitative study to identify barriers to providing precision medicines for non-small cell lung cancer. Semi-structured telephone interviews were conducted with clinicians, test providers, and service commissioners providing precision medicines to patients with NSCLC. The results of this study were used to identify the relevant capacity constraints to include in the case study decision analytic model.
Chapter 5 aimed to replicate a published decision-analytic model-based cost-effectiveness analysis of an example of precision medicine in lung cancer. The replica decision-analytic model was created to provide baseline estimates of the cost-effectiveness of a test-treat strategy for NSCLC in the absence of health system capacity constraints. The model represents a reproduction of the model submitted by Pfizer to NICE for the technology appraisal of crizotinib for ALK mutation positive patients. Additional detail regarding the testing pathway for ALK mutations was also added to the model. The predicted outcomes of the model were compared to those estimated in the manufacturer submission to determine the validity of the case study model.

Chapter 6 aimed to quantify the impact of selected capacity constraints on the relative cost-effectiveness of an example of precision medicine in NSCLC. A selection of capacity constraints identified in chapter 4 were incorporated into the model created in chapter 5 to determine the impact on the expected cost-effectiveness and value of the intervention at the time of evaluation. The adapted static value of implementation method developed in chapter 3 was used to quantify the impact of each capacity constraint and all of the constraints combined by determining the total net monetary benefit lost due the constraints.

Chapter 7 aimed to investigate the value of a defined number of capacity-investments to improve the ability of a constrained health system to provide ALK testing and targeted treatment with crizotinib to patients with NSCLC. Three hypothetical capacity investments aimed at reducing the impact of each capacity constraint were defined. The model created in chapter 5 and altered in chapter 6 was adapted to become dynamic, with changing patient populations and natural diffusion of the intervention. The dynamic value of implementation approach allowing for varying marginal costs and benefits which was developed in chapter 3 was used to estimate the value of the capacity investments.

Chapter 8 aim to discuss the collated findings from the five empirical chapters in terms of the proposed method for accounting for capacity constraints, the limitations of the methods, and the implications of the findings of these studies for policy making and research.
Chapter 2

Accounting for Capacity Constraints in Economic Evaluations of Precision Medicine: a Systematic Review

2.1 Chapter 2 Overview

This chapter reports a published systematic review of systematic reviews of economic evaluations of precision medicine: Wright SJ, Newman WG, Payne K. Accounting for Capacity Constraints in Economic Evaluations of Precision Medicine: A Systematic Review. Pharmacoeconomics. 2019;37(8); p1011-1027 (see Appendix 2.1)

The aim of this systematic review was to identify if, and how, published economic evaluations of precision medicine had recognised and quantified the potential impact of health system capacity constraints. The motivation for this study was to establish what is currently known about the topic and to inform the development of a method to account for capacity constraints in later chapters of this thesis. The accepted version of Wright et al (2019) is now reproduced in chapter 2.
2.2 Abstract

2.2.1 Background

Precision (stratified or personalised) medicine is underpinned by the premise that it is feasible to identify known heterogeneity using a specific test or algorithm in patient populations and use this information to guide patient care to improve health and wellbeing. This study aimed to understand if, and how, previous economic evaluations of precision medicine had taken account of the impact of capacity constraints.

2.2.2 Methods

A meta-review was identified published systematic reviews of economic evaluations of precision medicine (test-treat interventions) and individual studies included in these reviews. Due to the volume of studies identified, a sample of papers published from 2007 to 2015 was collated. A narrative analysis identified whether potential capacity constraints were discussed qualitatively in the studies and, if relevant, which quantitative methods were used.

2.2.3 Results

A total of 45 systematic reviews of economic evaluations of precision medicine were identified, from which 222 studies focussing on test-treat interventions, published between 2007 and 2015, were extracted. Of these studies, 33 (15%) included a qualitatively discussed the potential impact of capacity constraints including: budget constraints; quality of tests and the testing process; ease of use of tests in clinical practice; and decision uncertainty. Quantitative methods (9 studies) to account for capacity constraints included: static methods such as capturing inefficiencies in trials or models and sensitivity analysis around model parameters; and dynamic methods which allow the impact of capacity constraints on cost-effectiveness to change over time.

2.2.4 Conclusions

Understanding the cost-effectiveness of precision medicine is necessary evidence but not sufficient evidence for its successful implementation. There are currently few examples of evaluations that have quantified the impact of capacity constraints, which suggests an area of focus for future research.
2.3 Introduction

Precision (stratified or personalised) medicine is underpinned by the premise that it is feasible to identify known heterogeneity using a specific test or algorithm in patient populations to guide patient care to improve health and wellbeing (95). There is no unified definition of precision (stratified or personalised) medicine but in practice to date it refers to using some test-treat combination to target an intervention (95,96). A variety of mechanisms are undergoing development to identify such heterogeneity in the outcomes of interventions and progression of disease in populations of patients informed by genomic, proteomic, transcriptomic and metabolomic strategies (96). The ability to determine which patients might be more likely to benefit from a treatment, avoid harmful side effects or experience more severe disease has driven the theoretical arguments that precision medicine is a cost-effective use of healthcare resources (13,97).

Determining the incremental cost-effectiveness of the strategies to deliver precision medicine (hereafter shortened to precision medicine) is important because diverting funding to such newer interventions will involve the reallocation of resources from other areas of medicine. The reallocation of funding may affect the health outcomes for relevant populations of patients; representing the opportunity cost of the new intervention. Economic evaluations provide a structured framework to provide evidence supporting whether the introduction of precision medicine is an effective use of healthcare resources.

There is some economic evidence supporting exemplars of precision medicine (98). However, even with such economic evidence, the introduction of precision medicine into health systems has been slower than anticipated potentially due to the volume of eligible patients. There is emerging anecdotal, and some empirical, evidence of factors limiting the uptake of precision medicine. In 2015, the charity Cancer Research UK (CRUK) published a report highlighting the significant delays in providing genetic mutation testing to patients (99) and estimated that in the previous year approximately 3500 patients may have missed out on receiving a medicine which may have improved their quality and length of life because of the absence of relevant mutation testing. Questions about the ability of the UK NHS to implement one of the case study examples of precision medicine, EGFR mutation testing and gefitinib, a specific EGFR inhibitor licenced for patients with EGFR mutation positive non-small cell lung cancer, without “substantial investment in time and resources” have been raised. It was believed that the NHS did have capacity to introduce such testing in the short timeframe (100). However, this proved not to be the case, with the CRUK report suggesting that only 52% of patients received EGFR mutation testing even four years after the approval of gefitinib by the National Institute for health and Care Excellence.
This case study introduces the potential for capacity constraints to be a key barrier to the introduction of precision medicine, even when these have been shown to be cost-effective in clinical practice.

No consensus definition of what constitutes a capacity constraint in the context of health care interventions exists in the literature. The Oxford English Dictionary (OED) defines capacity as ‘the amount that something can produce’ (85). This definition implies that it is necessary to be clear what is being produced. The capacity of a healthcare system could therefore be defined as its ability to produce some defined output. In keeping with the extra-welfarist normative underpinning, assumed by decision-making bodies such as NICE, the relevant output of a healthcare system has been defined as ‘health status’ measured using the QALY. The definition of a constraint is “something that controls what you do by keeping you within particular limits” (86). Combining the definitions of capacity and constraint can be used to propose a working definition of a capacity constraint in a healthcare system which has the goal of maximising health status: “Any factor which impedes or limits the amount of health status produced for a population of patients receiving specified interventions, or policies, provided by the healthcare system.”

The introduction of any healthcare intervention may be impeded by capacity constraints in a healthcare system, which may be particularly extensive and significant for examples of precision medicine due to their nature as complex interventions involving both a test and treatment element (87). There is some qualitative evidence describing the type of capacity constraints directly relevant to the uptake of precision medicine into practice, including a lack of laboratories providing tests, poor logistics resulting in slow test turnaround, a lack of training for clinicians, and insufficient funding for testing or treatments (91,101,102). The impact of such capacity constraints on the incremental cost-effectiveness of precision medicine has not been well described. This study aimed to identify a sample of published economic evaluations of precision medicine and describe if, and how, these economic evaluations had qualitatively discussed and quantitatively accounted for capacity constraints in the analysis.

2.4 Methods

This study used a two stage systematic review conducted and reported in accordance with published guidelines and reporting criteria (103). A published search of PubMed by Payne and colleagues suggested that there were a substantive number of previously published systematic reviews of economic evaluations of precision medicine and related areas such as personalised medicine, pharmacogenetics and pharmacogenomics (98). Therefore, a de novo systematic review that aimed to identify all previous economic evaluations of precision medicine would overlap
significantly with this previous body of work, requiring significant resources in terms of researcher time but unlikely to yield substantively different findings. This study therefore used a strategy to identify and collate a sample of published systematic reviews of economic evaluations of precision medicine. The identified sample of reviews were published between 2007, when the use of terms related to precision medicine began to occur regularly in the literature (96), and February 2017.

For the purpose of this review precision medicine was defined as an intervention that uses, for example, a test to “identify subgroups of patients with distinct mechanisms of disease, or particular responses to treatments” (95). Related areas included were therefore; precision medicine; stratified medicine; individualised medicine; genetic medicine; genomic medicine; personalised medicine; and targeted medicine (104). This systematic collation of systematic reviews published up to February 2017 was then used to identify an exemplar sample of individual economic evaluations and use this sample to identify if, and how, capacity constraints had been included in the published analyses. The review involved two stages. Stage one involved the systematic collation of a sample of systematic reviews of economic evaluations of precision medicine. Stage two involved creating a list of relevant individual economic evaluations of precision medicine.

2.4.1 Stage One

Table 2.1 summarises the inclusion criteria used to guide the relevance of published systematic reviews of economic evaluations of precision medicine. To be classified as a systematic review, the published study must have used a systematic approach to search databases of published literature with the aim of identifying all studies which addressed a specified research question.

<table>
<thead>
<tr>
<th>Aspect of Study</th>
<th>Inclusion Criteria</th>
</tr>
</thead>
<tbody>
<tr>
<td>Population</td>
<td>Any relevant group of patients</td>
</tr>
<tr>
<td>Intervention</td>
<td>A stratifying test, algorithm or test-treatment combination</td>
</tr>
<tr>
<td>Comparator</td>
<td>Current practice</td>
</tr>
<tr>
<td>Outcomes</td>
<td>Costs and consequences relevant to a full economic evaluation (CUA, CEA, CBA)</td>
</tr>
<tr>
<td>Study type</td>
<td>Systematic review</td>
</tr>
<tr>
<td>Availability</td>
<td>English: full text</td>
</tr>
</tbody>
</table>
2.4.2 Search Strategy

The MEDLINE (inception year: 1946) and EMBASE (inception year: 1980) databases were searched from database inception to February 2017, using an electronic search strategy, to identify all systematic reviews of economic evaluations of precision medicine as defined by this review. The electronic search strategy for this systematic review was based on a published economic evaluation search filter developed by the University Of York based Centre for Reviews and Dissemination (CRD) (105) and combined with terms relevant to precision medicine and a systematic review search filter, which were informed by a published strategy (98) (see Appendix 2.2). In addition, a selection of lead or senior (last) authors of published systematic reviews of economic evaluations of precision medicine were contacted by e-mail to determine whether they knew of other published systematic reviews not identified in the initial search relevant to the selected time period.

2.4.3 Selection Process

The abstracts identified in the electronic literature search were screened for relevance and inclusion in the review by a team of three reviewers at the Manchester Centre for Health Economics (SW, Hunter Moore and Sean Gavan). Each abstract was screened by two of the three reviewers. Disagreements on whether a study should be included were resolved by a fourth researcher (Niall Davison).

2.4.4 Data Extraction and Analysis

The number of systematic reviews of economics evaluations of precision medicine was recorded and their key details summarised using a table (Appendix 2.2). The identified systematic reviews were listed and then categorised into one of four categories depending on their stated focus of the systemic review of economic evaluations: 1) test and treat interventions across disease areas; 2) test only interventions across disease areas; 3) test and treat interventions within a given disease area; and 4) test only interventions within a given disease area. Details regarding the number of primary economic evaluations cited in each review were also recorded. These findings were then presented using a narrative summary.

2.4.5. Stage Two

Table 2.2 summarises the inclusion and exclusion criteria for the individual studies to define a sample of economic evaluations of test and treatment based precision medicine. Studies which
focussed on interventions which provide diagnostic information which has no impact on treatment were excluded from this study. A second restriction applied because of the volume of individual economic evaluations was limiting the identified studies to those published in the ten-years previous to the date of the search.

Table 2.2: Inclusion Criteria for Primary Economic Evaluations of Precision Medicine

<table>
<thead>
<tr>
<th>Aspect of Study</th>
<th>Inclusion Criteria</th>
</tr>
</thead>
<tbody>
<tr>
<td>Population</td>
<td>Any relevant group of patients</td>
</tr>
<tr>
<td>Intervention</td>
<td>A stratifying test or algorithm used to subsequently guide a specified treatment or type of treatments</td>
</tr>
<tr>
<td>Comparator</td>
<td>Current practice</td>
</tr>
<tr>
<td>Outcomes</td>
<td>Costs and consequences relevant to a full economic evaluation (CUA, CEA, CBA)</td>
</tr>
<tr>
<td>Study type</td>
<td>Primary economic evaluations (prospective or model-based)</td>
</tr>
<tr>
<td>Availability</td>
<td>English: full text</td>
</tr>
<tr>
<td>Time frame</td>
<td>Published after 2007 and up to February 2017</td>
</tr>
</tbody>
</table>

2.4.6 Search Strategy

A list of individual economic evaluations of precision medicine were identified from the reference lists of the systematic reviews identified in stage one. Grey literature studies which had been identified in the systematic reviews were retained if they met the inclusion criteria for individual studies. This involved a manual search which was facilitated using a database created in Excel 2010 (106).

2.4.7 Selection Process

To determine the relevance of the individual studies and inclusion of primary economic evaluations of test and treatment strategies, the abstracts of the identified studies were imported into the bespoke database in Excel and then double screened by two reviewers (SW and Martin Eden, Manchester Centre for Health Economics), with disagreements resolved by a third reviewer (KP).
To identify studies which had discussed capacity issues, a manual keyword (see Table 2.3) search of the PDF for each included study was conducted. To identify relevant keywords, a search of the title and abstracts of relevant theoretical papers was conducted using the Mendeley reference management software (107)(77,79,81,94,108–118).

### Table 2.3: Capacity Related Terms Used in Keyword Search

<table>
<thead>
<tr>
<th>Capacity</th>
<th>Barrier*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constrain*</td>
<td>Restrict*</td>
</tr>
<tr>
<td>Short (for short-run or short-term)</td>
<td>Implement*</td>
</tr>
<tr>
<td>Learn*</td>
<td>Inefficien*</td>
</tr>
<tr>
<td>Bottleneck</td>
<td>Scale</td>
</tr>
<tr>
<td>Utilis*</td>
<td>Utiliz*</td>
</tr>
</tbody>
</table>

#### 2.4.8 Data Extraction and Analysis

The individual studies were initially collated into one of four categories according to the aim of the parent systematic review (as outlined in the stage one data extraction). It was clear, however, that some systematic reviews had identified some studies as test-treat when they were test only strategies (and vice versa). Therefore, after review of the titles and abstracts of the individual studies, they were then reclassified into one of three categories: 1) economic evaluations of test and treatment interventions; 2) economic evaluations of test only interventions; and 3) studies which did not meet the eligibility criteria for this systematic review. Studies classified into the second and third categories were then excluded from this review.

The total number of papers identified and their characteristics were summarised (Appendix 2.4). Data were extracted by one reviewer (SW) using a data extraction table produced in Microsoft Word (119). Data extraction fields included: author; year; country; intervention and comparator; whether the study mentioned capacity and a brief extract where this was so; whether the study attempted to account for capacity constraints, and a brief description of the method to account for capacity constraints. Due to the large size of this study, the data extraction table provided only reported papers which as a minimum criterion discussed capacity issues in a qualitative manner (Appendix 2.7). Studies which attempt to quantitatively account for capacity constraints were summarised using the Consolidated Health Economic Evaluation Reporting Standards (CHEERS) statement (120). Themes in the discussion and analysis of capacity issues in the individual evaluations were identified using thematic analysis of the published text in the manuscript. These themes were then discussed in a narrative summary. The narrative summary described the capacity constraints identified in the literature, the extent of the problem of capacity in economic
evaluations, and the methods which have been used to deal these issues in economic evaluations of precision medicine.

2.5 Results

A total of 45 systematic reviews of economic evaluations of precision medicine published up to and including February 2017 were included in this review (Appendix 2.3). A PRISMA diagram showing the process of identifying the systematic reviews for inclusion in this study is presented (Figure 2.1). The initial literature search identified 3,304 potentially relevant papers in the MEDLINE database and 990 in the EMBASE database. Microsoft Excel was used to remove duplicate copies first by abstract (n=283) and secondly by title (n=92) leaving 3,919 papers to review. During double screening of the abstracts of the identified papers, 3,871 papers were removed; 3,242 were not systematic reviews, 329 did not focus on precision medicine, 133 were not in English, 105 were duplicates, and 62 were not economic evaluations. This left 48 systematic reviews of economic evaluations of precision medicine to which an additional single study was added following direct contact with key authors in the area. Three reviews were subsequently excluded from data extraction because they did not report the citations for the individual studies included in their review and one study was removed as closer inspection it became clear it was not a systematic review of economic evaluations.
A total of 222 individual economic evaluations of precision medicine involving a test-treat strategy published between 2007 and 2015 were included in this review (see figure 2.2). Extraction of the individual economic evaluations from the identified systematic reviews yielded a list of a total 1,101 studies. From this group 477 papers were duplicates and a further 258 papers were removed as they did not evaluate an intervention relevant to precision medicine as defined in this review. Of the remaining 366 studies, a further 62 studies that reported an economic evaluation of a test only strategy that did not inform a treatment option were excluded. Restricting the individual economic evaluations of test-treat strategies to those published within the last ten years (2007-2017) yielded 259 studies. Of these studies a further 37 studies were removed during data extraction as it was clear on reading the full text that the study did not focus on precision medicine. Such studies were economic evaluations of medicines which had, in theory, a precision application but were not being evaluated in this way because the costs and consequences of the ‘test’ element was not included in
the evaluation. For example, some studies evaluated the cost-effectiveness of erlotinib for all non-small cell lung cancer patients rather than just those with \textit{EGFR} mutations (121,122).

**Figure 2.2: Identification of Primary Economic Evaluations of Test and Treatment Precision Medicine**

2.5.1 Description of Systematic Reviews

A total of 45 systematic reviews of economic evaluations of precision medicine were identified which were published between 2004 and February 2017. There were four categories of reviews of economic evaluations; test and treat interventions across disease areas or technology (n=13); test and treat interventions in a specific disease area of technology (n=17); diagnostic only interventions across disease areas or technologies (n=9); and diagnostic only interventions in a specific disease or technology area (n=6). The size of the reviews ranged from zero included papers, for a review of evaluation of genetic diagnosis of aneuploidy in all chromosomes (123), to 140 papers from a review of economic evaluations of targeted and non-targeted therapies for breast cancer (124).
2.5.2 Description of Individual Economic Evaluations

Of the 222 identified economic evaluations of precision medicine, 159 (71%) used quality-adjusted life years (QALYS) as the outcome of interest and 45 (20%) used clinical outcomes. A further 19 (9%) studies used a mixture of outcomes. The majority of the economic evaluations were decision analytic model-based (n=203, 91%), and 20 (9%) were based on primary data from trials. Of the studies using model based evaluations, 124 (56%) used Markov models, 33 (15%) used decision trees, 17 (8%) used linked decision trees and Markov models, and 16 (7%) were individual level simulation models. Six studies used other methods ranging from simple quantitative calculations based on literature reviews and meta-analysis to more complex methods such as dynamic life-cycle modelling, a system dynamics based approach (48). In 10 (4%) studies it was unclear what modelling approach was used.

The individual economic evaluations identified the costs and consequences of precision medicine for 32 distinct conditions or groups of conditions (Appendix 2.6). In total precision medicine targeting cancer comprised 67% (n=151) of the individual studies. Evaluations of precision medicine targeting breast cancer dominated the identified economic evaluations (n=104, 46%). The other common conditions of focus were cardiovascular conditions (n=22, 10%), including acute myocardial infarction, atrial fibrillation and acute coronary syndrome. Collectively the interventions in these studies were commonly aimed at preventing strokes. Other more commonly considered conditions included lung cancer (n=18, 8%), colorectal cancer (n=15, 7%), and HIV (n=12, 5%).

2.5.3 Inclusion of Capacity Constraints

Of the 222 individual economic evaluations included in this review, 33 mentioned the potential impact of capacity constraints on the costs and consequences of precision medicine in a qualitative sense and nine of these 33 studies went further and attempted to quantify the impact of capacity constraints in the analysis (table 2.4). Appendix 2.7 summarises the 33 included studies. These 33 studies raised key elements related to capacity constraints and how these may impact on the cost-effectiveness of precision medicine. These capacity constraints were grouped into four themes: budget constraints; quality of the testing process; ease of test use in clinical practice; and the need for economic evidence to reduce decision uncertainty.

2.5.4 Budget Constraints
A key concern relevant to capacity constraints was the total impact of precision medicine on healthcare budgets (budget impact). Budget impact was mentioned in terms of the specific treatment or associated testing and how this may potentially inhibit the uptake and use of these interventions in clinical practice. Three studies focused on the budget impact in broad terms (125–127) but a further ten studies specified the discussion about the financial impact of testing or treatment (20,126,128–135).

Kondo et al. (2008) stressed the need for concern regarding the financial impact of implementing 12,000 new 21-gene assays for breast cancer per year (20). The considerably larger eligible patient populations who could benefit from interventions such as warfarin dosing tests was suggested to pose more substantial problems in terms of financial impact. Patrick et al., identified that up to 10 million people could benefit from tests costing between $400 and $550 per test (25).

When evaluating the implementation of the more advanced fluorescent in-situ hybridisation (FISH) testing for reflex testing human epidermal growth factor receptor 2 (HER2) amplifications in breast cancer, Garrison et al., stated that it was unclear whether payers would finance the more rigorous test over the commonly used immunohistochemistry test (IHC) (129). Retèl et al., assessed that there was a 75% chance that a 70-gene signature test for breast cancer would not be immediately reimbursed by payers thereby limiting the availability of testing to patients (136).

The introduction of genetic testing for breast cancer was seen, by the Medical Advisory Secretariat of the Ministry of Health and Long-Term Care, Ontario, as an action which could either release health system resources by reducing the number of women receiving unnecessary chemotherapy, or place an additional burden through additional test costs (134). This report highlighted potential economic incentives which may impact the level of usage of testing, including pharmaceutical companies’ desire to maintain profits in providing the drug to wider population and payers’ desire to reduce drug expenditure.

The cost of specific treatments was also commonly highlighted as a potential barrier to their full implementation. Ito et al., (2013) suggested that the out-of-pocket costs of accessing aromatase inhibitors for breast cancer were suggested to cause under-utilisation of the potentially health improving treatment in the US (130). It was determined that improving coverage of these drugs by Medicare, a government health insurance scheme, would improve health outcomes while reducing healthcare resource use.

Lidgren et al., (2008) suggested that individual clinicians may not decide to prescribe treatments due their high cost (132). The authors also suggested that if different healthcare clinics have
different financial budgets then provision of trastuzumab for breast cancer may be variable and this could lead inequitable access to treatment.

2.5.5 Quality of Tests and Testing Processes

A commonly mentioned issue linked to capacity constraints and the cost-effective use of precision medicine was the quality of the tests and testing process to determine the appropriateness of treatment. Sub-optimal testing may result from factors such as limited supply of trained test providers such as pathologists and geneticists, testing facilities, or unclear reporting of test results to prescribing clinicians. The reduced test quality and potentially volume of available tests may both serve to reduce the potential health that could be produced by a new precision medicine. In a health technology appraisal report, Collinson et al., outlined the criteria for an effective cardiac biomarker: “A marker suitable for routine clinical use must be measurable in the routine clinical laboratory without special handling conditions” (137). These authors stated that a biomarker must be measurable with “precision and accuracy, the analysis must be simple and have a rapid turnaround time”. The biomarker test should also ideally be able to be implemented using existing laboratory resources. Eight studies discussed how current testing methods may fail to exhibit such characteristics.

Garrison et al., (2013) suggested that it may not be possible to fully implement reflex FISH testing for HER2 mutations in breast cancer given the need for additional laboratory resources and expertise (129). These authors explicitly stated a need for “improvements in capacity and investment in equipment, which may be costly”. The Medical Advisory Secretariat, Ontario, suggested similar barriers when evaluating FISH testing for non-small cell lung cancer, in particular a lack of expertise in the technique (138). A similar study by Djalalov et al., (2014), based in Canada, highlighted issues with the complexity of testing and the need for sufficient tissue alongside the relatively small number of patients with relevant mutations in a significantly larger population (128).

Romanus et al., (2015), in a US based study, identified that EGFR testing for non-small cell lung cancer was only available for 5.7% of the patients for whom guidelines specify testing should be available (139). These authors went on to suggest that long turnaround times for the tests used to target the medicine could push the balance of cost-effectiveness in favour of generic chemotherapy for all patients. Lala et al (2013) reported that shorter turnaround times for testing were also seen as a factor in making the implementation of point-of-care biomarker testing for adverse cardiovascular events feasible (140). In a survey conducted to input into the analysis of an economic evaluation, Barone et al., (2014), an Italian research team, found that 75% of a sample of
oncologists, pathologists, molecular biologists and surgeons believed a delay in receiving KRAS mutation testing result for colorectal cancer patients would affect their treatment choice and 25% experienced such delays in actual practice (141).

### 2.5.6 Ease of Test Use in Clinical Practice

To effectively introduce a new test into clinical practice, the relevant health professionals must be aware of the test and have sufficient training to offer, interpret and appropriately use the results in their decision-making to guide clinical practice. Insufficient human capital can therefore act as a capacity constraint in moving precision medicine into practice. When tests and treatments are assessed in trials and economic evaluation the assumption is that they are used in an optimal, or near optimal way, and with full adherence to their recommended use by clinicians. However, clinical practice is complex and the way in which a test fits into the current care pathway is not always clear. Nine studies highlighted issues with the transition of tests from research to clinical settings. For example, Breijer et al. (2012) developed and evaluated two multivariable models to produce an algorithm to predict the risk of endometrial cancer based on patients’ characteristics (142). While this analysis suggested a reduction in the cost of diagnosing the cancer, these authors proposed that due to the minimal expected value per patient, the algorithm would have to be made very easy to use by clinicians for it to be implemented in clinical practice.

The potential reluctance of clinicians to change their prescribing behaviour was shown in a UK study of TMPT genotyping to inform the use of azathioprine and the appropriate dosing strategy (15). Even though the test could be used to better predict which patients would experience profound neutropenia as a reaction from azathioprine, clinicians did not use the test to change the dosing of azathioprine. The authors hypothesised that this is because the clinicians chose to remain conservative with azathioprine prescription due to the other potential adverse reactions that the medicine could cause.

Retèl et al., (2012) predicted that the use of a 70-gene signature in breast cancer patients would be delayed by “hesitant adopters” who would not use the results of the assay in decision making. However, these authors included a scenario in their analysis, which assumed an increase in the ease of use of the 70-gene signature would subsequently increase the use of test results in clinical decision making resulting in improved cost-effectiveness of the intervention (136).

Two studies suggested that the willingness of clinicians to use new tests and their ability to effectively use results in decision making may evolve over time as a result of a learning process (143,144). Klang et al., (2010) found that clinicians using the Oncotype DX panel did not register
treatment decisions for the first 55 of 368 patients as the clinicians were “learning about the technology and how to interpret the results” (143). The authors of a US study investigating the use of a multi-gene assay in breast cancer explicitly highlighted that previous analyses of Oncotype Dx were based on the assumption that the test would be used as dictated by guidelines (145). They further stated they believe policy makers were interested in “learning how the assay affected outcomes and costs compared with actual practice and after some period of experience with the assay”. Both of these studies indicated that a lower utilisation of testing by health professionals may impact on the incremental cost-effectiveness of precision medicine in clinical practice.

### 2.5.7 Decision Uncertainty

When a new precision medicine is introduced, decision makers, such as individual clinicians, local hospital trusts or national HTA agencies, must decide under conditions of uncertainty whether to provide the intervention to patients (implementation). There is a significant cost associated with making the wrong decision associated with implementation which could include reduced patient outcomes at an individual level or societal loss of health from funding cost-ineffective interventions. While commonly used methods such as deterministic and probabilistic sensitivity analysis allow for uncertainty in model parameters to be visualised, a wide range of other uncertainties such as methodological and structural uncertainty may be present in the evaluation of precision medicine (146,147). Decision uncertainty around implementation can therefore act as a significant capacity constraint to the introduction of new examples of precision medicine.

As a rationale for conducting economic evaluations, many studies highlighted that evidence of the cost-effectiveness of an intervention was a requirement for implementation in a clinical setting (148,149). As such, a lack of such evidence on incremental cost-effectiveness could inhibit the use of precision medicine in clinical practice. In Canada, the Medical Advisory Secretariat, Ontario, suggested that evidence on the net resource implications of HER2 testing for breast cancer was required before the intervention could be adopted into “dynamic health systems” (134).

The published evidence also recognised that economic evaluations are conducted at a certain stage in the development of the intervention to deliver precision medicine and that the use and cost-effectiveness may change in subsequent clinical practice. Retèl et al. (2010), highlighted that their results were based on the assumption of full implementation and that an evaluation using real life scenarios was in progress at the time of publishing this study (136,150).

In theory, the availability of economic evidence can help to fuel increased implementation of precision medicine but in addition, increased implementation can itself provide economic benefits.
Rubinstein et al. (2009) identified that if decision makers relied on a passive diffusion of \textit{BRCA} 1/2 genetic testing into the healthcare system then potentially the economies of scale which could have been generated from a more managed implementation strategy could be forgone (151). These authors also suggested that increased implementation of an alternative intervention magnetic resonance imaging in breast cancer, may create additional costs and benefits if the imaging becomes a part of the cancer management pathway alongside \textit{BRCA} 1/2 genetic testing. This synergistic effect is known as an economy of scope whereby the same resource can be used to produce or improve multiple services to provide better outcomes.

\textbf{2.5.8 Quantifying the Impact of Capacity Constraints in Economic Evaluations}

Nine (see Table 2.4) of the identified 222 economic evaluations of precision medicine included in this systematic review used, or suggested, techniques that explicitly quantified the impact of capacity constraints. All of these studies used decision-analytic models that allowed for sub-perfect implementation of technologies, limiting their potential benefit to society. In other words, for various reasons, fewer than 100\% of the eligible patient population were assumed to receive the intervention delivering precision medicine or the intervention was not given in the optimal way resulting in higher costs or lower benefits than were potentially achievable. In some cases this less than perfect implementation was due to capacity issues such as budget constraints (136), regulation barriers (136), and long test turnaround times (139). In two other instances, the imperfect implementation was associated with low uptake of the test or treatment (152,153). While adherence to medication may appear to be a demand side problem, it is in fact a complex issue and “the attributes of the health-care system and service delivery may also influence adherence” (153). For example, if clinicians have limited time to spend with patients, the opportunity to provide effective information about the benefits of adherence and approaches to coping with side effects may be limited. The methods used to account for capacity constraints can be categorised as static or dynamic and are now described.
Table 2.4: Summary of included studies using CHEERs checklist

<table>
<thead>
<tr>
<th>Author (Year) Country</th>
<th>Intervention and comparator</th>
<th>Study Population</th>
<th>Economic evaluation type</th>
<th>Evaluation vehicle (model type if applicable)</th>
<th>Time Horizon (Discount Rate)</th>
<th>Analysis</th>
<th>Approach to quantifying capacity constraints</th>
</tr>
</thead>
<tbody>
<tr>
<td>Delea et al., (2012) UK</td>
<td>Intervention: Lapatinib and capecitabine Comparator: Capecitabine monotherapy</td>
<td>Women with HER2 positive metastatic breast cancer who have previously received trastuzumab</td>
<td>Cost-utility analysis</td>
<td>Model (partitioned survival analysis)</td>
<td>5 years (3.5%)</td>
<td>Incremental analysis reported: Yes PSA: Yes Other sensitivity analysis: One-way deterministic</td>
<td>An average drug wastage number was used in the analysis and this was set to zero in sensitivity analysis, reducing total costs. This suggests cost-effectiveness of intervention depends on implementation.</td>
</tr>
<tr>
<td>Delea et al., (2013) UK</td>
<td>Intervention: Lapatinib and letrozole Comparator: Trastuzumab and anastrozole or</td>
<td>Women with hormone receptor and HER2 positive metastatic breast cancer</td>
<td>Cost-utility analysis</td>
<td>Model (partitioned survival analysis)</td>
<td>10 years (3.5%)</td>
<td>Incremental analysis reported: Yes PSA: Yes Other sensitivity analysis: One-way deterministic</td>
<td>An average drug wastage number was used in the analysis and this was set to zero in sensitivity analysis, reducing total costs. This suggests cost-effectiveness</td>
</tr>
<tr>
<td>Study</td>
<td>Intervention</td>
<td>Comparator</td>
<td>Model</td>
<td>Lifetime Analysis</td>
<td>Incremental Analysis</td>
<td>Sensitivity Analysis</td>
<td>Comments</td>
</tr>
<tr>
<td>----------------------</td>
<td>---------------------------------------------------</td>
<td>-------------------------------------------------</td>
<td>------------------------------------</td>
<td>------------------</td>
<td>----------------------</td>
<td>-----------------------</td>
<td>--------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Djalalov et al., (2014) Canada</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
| Intervention: *EML4-ALK*  
Fusion testing and first-line crizotinib treatment |
<p>| Comparator: Cisplatin and gemcitabine |
| Patients with advanced <em>ALK</em>-positive non–small-cell lung cancer |
| Cost-utility analysis |
| Model (Decision tree linked to Markov Model) |
| Lifetime (5%) |
| Incremental analysis reported: Yes |
| PSA: No |
| Other sensitivity analysis: One-way and two-way deterministic |
| Decision tree includes a branch for whether there is an adequate tissue sample and if not allows for a second biopsy to be taken. It is not clear if these probabilities were varied in sensitivity analysis but the cost of re-biopsy was allowed to vary. |
| Garrison and Veenstra (2009) USA |
| Intervention: Trastuzumab |
| Comparator: Not stated |
| Women with various stages of breast cancer |
| Cost-utility analysis |
| Model (dynamic life-cycle modelling) |
| 10 year product lifecycle (3%) |
| Incremental analysis reported: Yes |
| PSA: No |
| Other sensitivity analysis: One-way deterministic |
| Dynamic cost-effectiveness with changing patient population. Implies limited approved indications for drug may inhibit potential cost-effectiveness |</p>
<table>
<thead>
<tr>
<th>Lorenzana et al., (2012)</th>
<th>South Africa</th>
<th>Intervention: genotype assay for selection of third-line antiretroviral therapy (ART)</th>
<th>ART naïve cohort of patients with HIV</th>
<th>Cost-effectiveness analysis</th>
<th>Model (Discrete Event Simulation)</th>
<th>Time horizon not stated (3%)</th>
<th>Incremental analysis reported: Yes</th>
<th>PSA: No</th>
<th>Other sensitivity analysis: One-way and multi-way deterministic</th>
<th>Test cost was varied in sensitivity analysis with suggestions that higher test cost could represent cost when investment is accounted for. No impact on cost-effectiveness found.</th>
</tr>
</thead>
<tbody>
<tr>
<td>McCowan et al., (2013)</td>
<td>UK</td>
<td>Intervention: High adherence (≥80%) to tamoxifen</td>
<td>Women with breast cancer</td>
<td>Cost-utility</td>
<td>Model (Markov)</td>
<td>Lifetime (3.5%)</td>
<td>Incremental analysis reported: Yes</td>
<td>PSA: Yes</td>
<td>Other sensitivity analysis: One-way deterministic</td>
<td>Evaluation conducted across sub-groups of patients with under or over 80% adherence. Low adherence associated with expected loss of 1.12 discounted QALYs and</td>
</tr>
</tbody>
</table>
In Retèl et al., (2012) study in The Netherlands, the intervention involved a 70-gene Mammaprint assay to guide adjuvant breast cancer treatment. The comparator was the Adjuvant! Online algorithm to guide treatment. Women with breast cancer were the subject group. The model used was a cost-utility model, which includes a linked decision tree and Markov model with multiple cohorts and varying parameters, spanning 15 years with 4% costs and 1.5% outcomes. Incremental analysis was reported as yes, with no PSA analysis. Other sensitivity analysis was not conducted. The researchers modelled the cost-effectiveness over time and diffusion of the technology, considering a range of potential scenarios and barriers which affect the diffusion of the technology.

In Romanus et al., (2015) study in the USA, the intervention was multiplexed testing for EGFR and ALK in patients with NSCLC. The model used was a cost-utility model with a discrete event simulation, spanning 2 years with 3% costs. Incremental analysis was reported as yes, with no PSA analysis. Other sensitivity analysis was not conducted. The researchers included a parameter for turnaround time and inadequate tissue sample leading to re-biopsy as well.
<table>
<thead>
<tr>
<th>Vanderlaan et al., (2011) USA</th>
<th>Intervention: 21 gene assay to guide adjuvant chemotherapy Comparator: treatment guided by US NCCN guidelines</th>
<th>Women with node-positive, early stage breast cancer</th>
<th>Cost-utility</th>
<th>Model (decision tree)</th>
<th>30 years (3%)</th>
<th>Incremental analysis reported: Yes PSA: No Other sensitivity analysis: One way deterministic.</th>
<th>Sensitivity analysis included variations in utilisation rates of testing, although marginal costs were linear so no impact on cost-effectiveness</th>
</tr>
</thead>
<tbody>
<tr>
<td>mutations to guide NSCLC treatment Comparator: No testing and treatment with pemetrexed and cisplatin</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Other sensitivity analysis: One way deterministic. Threshold analysis for turnaround time for testing</td>
<td>as proportion of patients tested</td>
</tr>
</tbody>
</table>
2.5.9 Static Methods to Account for Capacity Constraints

Static methods refer to methods used to produce a single cost-effectiveness estimate which takes account of imperfect implementation for one cohort of patients. This estimate of the incremental cost-effectiveness may differ from the estimated incremental costs and consequences for a perfectly implemented precision medicine. For example, Delea et al., (2012 and 2013) accounted for drug wastage in their model-based economic evaluation of lapatinib and letrozole for women with HER2 positive breast cancer (154,155). This analysis identified that reducing wastage of trastuzumab from 15% to 0% resulted in lapatinib and letrozole being cost-effective, at £22,895 per QALY gained when compared with trastuzumab, which was cost-saving.

Romanus et al., (2015) accounted for specific capacity constraints to multiplexed biomarker testing for non-small cell lung cancer in a sensitivity analysis (139). When turnaround time increased by a factor of 1.5, the most cost-effective approach changed from “test and treat” to “empiric therapy”, in which patients began treatment with a general chemotherapy agent while waiting for test results. Reducing the proportion of patients being tested from 100% to 5.7% did not impact the rank ordering of the incremental cost-effectiveness of the interventions. Likewise, in a study by Vanderlaan and colleagues (2011) assumptions about the differences in the uptake of a 21-gene assay for breast cancer did not impact on the incremental cost-effectiveness of the intervention (152). In two further studies, assumptions about the cost of the test or re-biopsy did not affect the incremental cost-effectiveness of a genotype assay for drug selection in patients with HIV (133) or EML4-ALK testing for patients with non-small cell lung cancer (128).

McCowan et al., (2013) found that adherence to tamoxifen in women with breast cancer was found to significantly impact the incremental cost-effectiveness (153). On average, patients with an adherence of less than 80% to the treatment were expected to experience 1.12 fewer QALYs compared with those patients assumed to have over 80% adherence. Furthermore, such patients were expected to experience significantly higher medical costs (£5970, 95% CI: £4,644 to £7,372).

2.5.10 Dynamic Methods to Account for Capacity Constraints

Dynamic methods refer to methods that account for capacity constraints which allow the impact of barriers or constraints to change over time and/or in multiple patient cohorts. As a result, the cost-effectiveness of the technology also potentially changes over time. The most comprehensive
investigation of capacity constraints and impact on implementation in an economic evaluation using dynamic methods was conducted by Retèl and colleagues who investigated the future potential uptake of a 70-gene signature test in breast cancer (136). This study used an analytical approach called scenario drafting to identify potential barriers and facilitators to the implementation of the technology. These authors constructed different sets of model parameters to reflect the implications of these scenarios, accounting for their perceived likelihood of occurring in reality. These scenarios comprised potential capacity issues such as a lack of reimbursement of testing, regulation issues, uncertainty in the clinical utility of the test, and a lack of use by clinicians due to the difficulty interpreting the tests. The authors also investigated the potential impact of using the test at different stages of the clinical pathway.

The incremental cost-effectiveness of the 70-gene signature test was evaluated at different time points and implementation levels for three key scenarios: reducing technical failure rates of the test over time, decreasing non-compliance with discordant test results, and increasing financial reimbursement and clinicians’ uptake of testing. The cost and consequences achieved by testing an additional patient was allowed to vary. As each scenario involved the intervention diffusing into clinical practice at different rates, the costs, health outcomes, and cost-effectiveness of the intervention differed by year and scenario. A key finding of the study was that the intervention would only be cost-effective if use of the test results by clinicians improved over time. At the initial time point, 2005, the intervention was always cost-ineffective with an incremental cost-effectiveness ratio (ICER) of €1.9million. If uptake did not improve then the ICER only improved to €1.5million in the best case scenario. In the scenario where other factors remained equal and uptake improve from 3% in 2005 to 50% in 2010 and 92% in 2010, the ICER reduce from €1.9million to €26,145 and €11,123 respectively. Failing to overcome the capacity constraints of low financial reimbursement of the test and a lack of uptake by clinicians would therefore mean a potentially cost-effective intervention should not be provided as it would result in a societal loss of health to the population. Removal of the capacity constraints makes the intervention cost-effective, but the cost of such capacity investments would also need to be accounted for in an economic evaluation.

2.6 Discussion

Precision medicine is often reported to offer a cost-effective approach for the management of a selection of different diseases by using a stratification mechanism to identify which patients may
accrue more benefits in terms of response or avoidance of adverse events. Some economic evidence supporting this premise is provided by the growing number of economic evaluations conducted in the past ten years. This review identified 45 systematic reviews of economic evaluations of precision medicine which summarised 367 unique individual economic evaluations. Between 2007 and 2015, some 222 economic evaluations were conducted to identify the costs and consequences of test and treat interventions.

The first necessary condition for the adoption of an example of precision medicine is demonstration whether it is a cost-effective use of healthcare resources using appropriate methods of economic evaluation to identify incremental costs and consequences. However, capacity constraints in a healthcare system have the potential to impact on the estimated incremental cost-effectiveness of examples of precision medicine in practice. Only 33 studies of a sample of 222 identified economic evaluations qualitatively discussed the potential impact of capacity constraints for the introduction of examples of precision medicine. The core capacity constraints outlined in these papers included: budget constraints; quality of tests and testing processes; ease of test use in clinical practice; and decision uncertainty. Even when interventions appeared cost-effective, it was suggested they can pose significant financial burdens on payers who have to rapidly provide access to testing services and potentially expensive treatments.

Despite the potential for capacity constraints to affect the incremental cost-effectiveness of precision medicine, only nine (4%) economic evaluations from the sample of 222 sought to quantify the effects of limited capacity. All of the nine studies which quantified the effects of limited capacity were based on decision analytic models rather than trials and a wide range of model types including decision trees, Markov models and discrete event simulation were used. The presence of these commonly used models in this review suggests that it would be possible to incorporate capacity constraints in many economic evaluations.

While the way in which capacity constraints are included in models will depend on the model type, the methods used could be categorised as either a static or dynamic approach. The static methods included real-world cost-effectiveness analysis and sensitivity analysis in decision analytic models. Static approaches to incorporating capacity constraints only give a representation of how the constraints impact the cost-effectiveness of the intervention for a single cohort of patients at a single point in time. Dynamic approaches to quantifying capacity constraints allow for the fact that health system capacity can change over time and this can have a changing impact on the cost-
effectiveness of the intervention. The example of Retèl et al focussed on how the level of uptake impacts on the cost-effectiveness of the intervention (136). However, the method used could also be applied to capacity constraints. If the ICER is non-linear and depends on the level of implementation, then any factor which impedes implementation can in theory render the intervention cost-ineffective. For example, if numerous repeat EGFR mutation tests are required due to insufficient samples at the start of implementation, this will raise the cost of testing and reduce patient benefits due to delays in receiving treatment. With greater communication regarding biopsy requirements and learning by clinicians, better samples could be obtained, reducing costs and improving benefits. Therefore, the ICER will be dependent on the extent to which the intervention is being effectively implemented and the impact of testing knowledge as a capacity constraint. This study shows the key interaction between barriers that impede the use of an intervention and the potential for marginal costs and benefits to vary depending on the level of implementation. The result is situations where capacity constraints and other barriers to implementation cause the intervention to become cost-ineffective. Implementing the precision medicine at this level will result in healthcare resources being diverted away from other areas where they could be put to better use, lowering the overall health of patients in the healthcare system.

The Retèl et al., study also raised an important consideration in that actions can be taken to overcome capacity constraints and to improve the incremental cost-effectiveness of precision medicine. For example, testing guidelines could be introduced or education programmes used to improve the quality of biopsy samples for testing. Such an intervention itself would have a cost but would also provide benefits in improving the benefits for patients and making the intervention more cost-effective. The evaluation of such strategies to improve the implementation of interventions is known as value of implementation analysis (114). This systematic review did not identify any value of implementation analyses investigating the implementation of precision medicine.

The use of dynamic, multi-cohort decision analytic models and the value of an implementation approach may provide better evidence than static methods to decision makers in allowing them to understand how best to implement new examples of precision medicine in a cost-effective way by investing in improving health system capacity. In order to forecast the potential capacity constraints before the approval of a precision medicine by an HTA body, the use of qualitative methods such as interviews or focus groups may prove useful. For example, Retèl et al used the
The Delphi method with a group of clinicians to identify potential barriers to implementing the Mammaprint test (136).

The economic evaluation of precision medicine is an expanding research area and due to the size of the literature base, some restrictions to the inclusion of papers in this study were required. The use of a search strategy which identified previous systematic reviews reduced the number of papers for abstract screening. Collating these studies provided a comprehensive set of primary studies and theoretically all papers which had previously been published in this area. However, due to the length of time required to conduct a systematic review, it is possible that economic evaluations of precision medicine have been published in the time since the literature searches of the most recent systematic reviews were conducted. This means there may be a gap in the studies identified between 2015 and 2017. This could cause bias in the results of this study if methods to include capacity constraints in economic evaluations have been recently developed and applied. The authors of this review are not aware of any recently novel published methods for accounting for capacity constraints in economic evaluations of precision medicine.

In addition to using a restricted sample of primary economic evaluations, a key word search was used to facilitate the identification studies which potentially discussed or accounted for health system capacity constraints. While these search terms were taken from studies discussing the theoretical impact of capacity constraints on economic evaluations, the results of the review suggested that there were potential additional terms which would have been relevant to include in the initial search. For example, terms related to the quality of testing, the need to wait for test results, and the budget impact of tests or treatments may have yielded additional relevant studies. For future systematic reviews using key words searches to identify relevant studies, an iterative approach whereby new key words are added as they are identified in the included studies may be useful.

The focus on studies published in the last ten years and evaluations which focussed on test and treatment strategies could also feasibly have excluded studies which discussed capacity. While studies published before 2007 may have discussed capacity due to the novelty of precision medicine, it is unlikely that early economic evaluations of precision treatments would have incorporated complex methodological adjustments for capacity. Capacity could be a significant issue in the provision of precision treatments which only provide diagnostic information. Such studies were excluded from this review.
This review has taken a broad definition of a capacity constraint as any factor which impedes the full benefits of an intervention from being realised. While this includes factors such as finite budgets and the quality of the testing process, it also includes more abstract concepts such as low usage of tests due to a lack of knowledge of the technology amongst clinicians. Some of the included studies also investigated adherence to medicine and uptake of treatment. While low adherence could be due to capacity constraints in a lack of patient education about the benefits of treatment, it may also be due to underlying patient preferences for the treatment or side effect profile. Therefore while full implementation of a precision medicine will rely on a lack of capacity constraints on the supply side, it could also be impeded by low demand for the treatment by patients. This could be the case if a new treatment had a greater risk of more side effects, more severe side effects, or a different range of side effects. There may, therefore, be a limit to the level of implementation that can be achieved by investing in capacity.

While this review has focussed on the impact of capacity constraints for the economic evaluation of precision medicine, such constraints in the healthcare system may have a significant impact on the cost-effectiveness of interventions in other medical areas. For example, Jahn et al., (2010) explored the cost-effectiveness of drug-eluting stents in the presence of capacity constraints using a discrete event simulation model (115). Capacity constraints may also be particularly significant for the cost-effectiveness of organ transplants where there is a limited availability of donors (156). In addition, capacity constraints and their impact may be larger in countries with developing healthcare systems (157).

2.7 Conclusions

The results of this systematic review suggest that a wide variety of capacity constraints could have implications for the cost-effectiveness of precision medicine in clinical practice. A variety of methods had been used to account for such constraints in economic evaluations, with most relying on static comparisons of the cost-effectiveness of examples of precision medicine for a single cohort with or without health system capacity constraints.

Studies should account for changing health system capacity which may have implications for the cost-effectiveness of interventions across multiple cohorts in different years. This is because when combined with varying marginal costs and benefits, implementation limiting capacity constraints
can result in interventions which are cost-effective at the population level becoming inefficient in the short-run. Health economists should endeavour to forecast potential barriers to implementing precision medicine and to evaluate potential strategies to invest in capacity. A number of studies have reported qualitative investigations of such barriers (91,101,102) and combining these approaches with dynamic methods for quantifying such barriers may help to provide decision makers with more robust evidence as to how to cost-effectively implement such interventions and take resource and capacity constraints into account.

2.8 Summary of Chapter 2

The meta-review reported in chapter 2 has shown that published economic evaluations of examples of precision medicine rarely account for health system capacity constraints. This may be a key omission in precision medicine given that these complex interventions may face a number of barriers to their introduction in clinical practice. Furthermore, this review has shown that when such capacity constraints have been incorporated into economic evaluations of examples of precision medicine they may impact the cost-effectiveness of the intervention. Given this potential impact, there may also be value in investing to reduce the impact of the capacity constraints. Chapter 3 will introduce a method, static and dynamic value of implementation, which can be used to value changes in the implementation of an intervention.
Chapter 3

Implementing Interventions with Varying Marginal Cost-Effectiveness: An Application in Precision Medicine

3.1 Chapter 3 Overview

The results of the systematic review described in chapter 2 indicated that it is possible that capacity constraints will influence the estimated cost-effectiveness of precision medicine. Evidence is therefore needed about the impact of capacity constraints on the cost-effectiveness of examples of precision medicine and appropriate methods to provide this evidence are required. Chapter 3 reports an empirical study currently under review with Medical Decision Making: Wright SJ, Paulden M, Payne K. Implementing Interventions with Varying Marginal Cost-Effectiveness: An Application in Precision Medicine.

The aim of this study was to describe the application of VOImp, using the static and dynamic forms to illustrate the potential impact of accounting for varying marginal costs and benefits on the relative cost-effectiveness and value of an exemplar of precision medicine. The motivation for this study was to develop a method which could quantify the impact of capacity constraints at different periods of time and value potential investments in overcoming these constraints. The version of Wright et al (under review) is now reproduced in chapter 3.
3.2 Abstract

3.2.1 Purpose

A range of barriers may constrain the effective implementation of strategies to deliver precision medicine. If the marginal costs and consequences of precision medicine vary at different levels of implementation, then such variation will have an impact on relative cost-effectiveness. This study aimed to illustrate the importance, and quantify the impact, of varying marginal costs and benefits on the value of implementation for a case study in precision medicine.

3.2.2 Methods

An existing method to calculate the value of implementation was adapted to allow marginal costs and consequences of introducing precision medicine into practice to vary across differing levels of implementation. This illustrative analysis used a case study based on a published decision-analytic model-based cost-effectiveness analysis of a 70-gene recurrence score (Mammaprint) for breast cancer. The impact of allowing for varying costs and benefits for the value of the precision medicine and of implementation strategies was illustrated graphically and numerically in both static and dynamic forms.

3.2.3 Results

The increasing returns to scale exhibited by introducing this specific example of precision medicine mean that a minimum level of implementation (51%) is required for using the 70-gene recurrence score to be cost-effective at a defined threshold of €20,000 per quality adjusted life year. The observed variation in net monetary benefit implies that the value of implementation strategies was dependent on the initial and ending levels of implementation in addition to the magnitude of the increase in patients receiving the 70-gene recurrence score. In dynamic models, incremental losses caused by low implementation accrue over time unless implementation is improved.

3.2.4 Conclusions
Poor implementation of approaches to deliver precision medicine, identified to be cost-effective using decision-analytic model-based cost-effectiveness analysis, can have a significant economic impact on health systems. Developing and evaluating the economic impact of strategies to improve the implementation of precision medicine will potentially realise the more cost-effective use of health care budgets.

3.3 Introduction

Precision (or stratified or personalised) medicine (96) has been defined as “an emerging approach for disease treatment and prevention that takes into account individual variability in “genes, environment, and lifestyle for each person” (1). In practice, precision medicine is underpinned by the premise that it is feasible to identify known heterogeneity within a disease or population and use this information to guide management strategies to improve health and well-being. Precision medicine, therefore, requires a mechanism (‘test’ or ‘tool’) that, in theory and in practice, provides information in addition to the currently available strategies used to select interventions, which can be for example “prognostic markers, predictors of toxicities and any parameter such as environmental and lifestyle factors” (2). These tools can be used to identify information about a patient and their disease to predict potential improved or reduced response to treatments, such as KRAS gene mutation testing to target cetuximab [3], or higher risk of side effects, such as CYP2C19 testing to guide the dose of warfarin to reduce major bleeding complications (7).

Precision medicine, in general, and the use of tests to inform the prescribing of medicines (test-treat interventions), specifically, have been suggested to have economic benefits by targeting interventions only to those patients who will accrue benefits and/or are less likely to experience severe adverse drug reactions (13).

Cost-effectiveness analysis can provide evidence about whether new healthcare interventions, such as test-treat interventions, represent a good investment by generating more health for patients receiving precision medicine compared with the health lost by those from whom funding is denied or removed (29). Decision-analytic model based cost-effectiveness analysis (hereafter ‘CEA’) is the cornerstone of the evidence-base in health technology assessment (HTA) reports produced as part of national decision-making processes by bodies such as the National Institute for Health and Care Excellence (NICE) (44) and the Canadian Agency for Drugs and Technologies in Health (CADTH) (158). A common, but not usually explicitly stated, assumption underpinning existing CEA is that marginal costs and benefits are constant (83,159). This assumption of constant
marginal changes implies that the mean incremental costs and incremental benefits of providing one more patient with the new intervention, such as a test-treat intervention, when compared with each comparator are the same regardless of the number of patients who are treated. The expected mean incremental cost and incremental benefit of the intervention for each patient is, therefore, assumed to represent the relative estimated cost-effectiveness of providing the test-treat intervention to all patients within a relevant population.

Existing applications of CEA also assume that the test-treat interventions under evaluation are divisible (64). Being divisible implies that it is possible to allocate precision medicine to a defined proportion of the relevant population. In practice HTA agencies tend to evaluate new test-treat interventions with the assumption that they are not divisible to promote equity. Actioning this assumption results in two distinct scenarios in which either all (or none) of the relevant population will receive the specified test-treat intervention (160). In some circumstances, the definition of the relevant population may be a clinically pre-specified sub-group of the population, for example a group of patients with a specific genetic mutation or a known level of disease severity. It has been argued that due to these assumptions, the cost-effectiveness estimates produced by conventional economic evaluations represent an estimate of the long-run cost-effectiveness of precision medicine (94,109). However, in the short-run these assumptions may not be realistic and the cost-effectiveness of precision medicine may differ as a result.

It is possible, but improbable, that the entire relevant population will have access to a test-treat intervention immediately following a positive HTA recommendation. A scenario of delayed, imperfect uptake is relevant to any intervention but precision medicine provides an exemplar of some specific challenges. A number of barriers (or capacity constraints) have been identified that may hinder fully comprehensive and inclusive access to the introduction of precision medicine: lack of sufficient laboratories to perform testing; logistical issues with coordinating testing and treatment; sufficient numbers of trained laboratory staff, allied healthcare professionals and clinicians (89,91,101,161). Concerns about the capacity of the National Health Service in England (NHS England) to provide the required EGFR mutation testing were raised during a NICE Technology Appraisal of gefitinib for NSCLC (100). Despite assurances at the time of the appraisal subsequent evidence in 2014 suggested that only around 50% of eligible patients were receiving EGFR testing (99). Issues such as geographical differences in the type of test being offered and long turnaround times for test results resulting in delayed treatment were a subsequent feature of NICE technology appraisals for erlotinib and afatanib (162,163). Such capacity
constraints potentially impede the comprehensive and inclusive evidence-based introduction (hereafter termed ‘implementation’) of precision medicine with its required combination of ‘test’ and subsequent ‘treatment’.

Methods are available to quantify the expected value of the improved implementation of healthcare interventions using specified strategies, such as investing in the required equipment or staff knowledge (164,165). Such approaches are henceforth referred to as ‘implementation-strategies’. The evaluation of implementation-strategies by estimating the incremental cost-effectiveness ratio was originally proposed by Sculpher in 2000 (117) and further developed by Mason et al., in 2001 (166). However, a move in 2007 to propose the use of the net benefit framework by Fenwick et al. (114), formulated by Walker et al. (167), allowed the calculation of the economic benefit of increasing implementation rather than a sole focus on quantifying the relative cost-effectiveness of specific implementation-strategies.

Underpinning these methods is the assumption that improving implementation results in a net health benefit from the intervention under evaluation. The gain in net health benefit of improvement in implementation is then compared with the cost of the implementation-strategy (activities of the health system to improve implementation) to quantify whether it is an appropriate use of healthcare system resources by estimating the value of implementation for a defined patient population and healthcare budget (VOImp). Two existing (‘traditional’) value of implementation frameworks (VOImp) are available that estimate static (VOImp (static)) and dynamic (VOImp (dynamic)) values (167). The static-approach can be used to calculate the value of using a one-time only implementation strategy to improve the implementation of a precision medicine that is cost-effective. The static-approach assumes that all the impact occurs in the first year. The dynamic-approach allows the costs and effects of the implementation strategy to improve implementation of a specific intervention across multiple time periods. In addition, methods exist to inform decisions as to whether to invest in implementation or additional research when faced with uncertainty in the parameters used to generate estimates of cost-effectiveness (168–170).

To date there have been few actual applications of VOImp, and each has focussed on estimating a ‘static’ value (164,165,171) and none have focussed on the impact on precision medicine. Whyte et al. (2016) estimated the value of using a two-hour training workshop for clinicians to increase the implementation of NICE guidelines for natriuretic peptide testing (164). Mewes et al. (2017) calculated the incremental net benefit of implementation activities to increase the use of novel oral
anticoagulants to prevent strokes (165). This study extended the VOImp framework to understand the value of implementation in different sub-groups of the patient population. Faria and colleagues estimated the VOImp of strategies to improve adherence to exercise guidelines for cancer survivors (171).

While VOImp methods can aid in decision making regarding the value of implementation strategies, to date no VOImp study has accounted for the potential presence of varying marginal costs and benefits which are likely to be a feature of many examples of precision medicine that rely on test-treat strategies. If the marginal costs and benefits of precision medicine vary then the cost-effectiveness and value of precision medicine will vary depending on the degree of implementation. This study aimed to illustrate the importance, and quantify the impact, of varying marginal costs and benefits on the value of implementation for a case study in precision medicine. The application of existing methods to calculate VOImp, using the static and dynamic forms, are applied to illustrate the potential impact of accounting for varying marginal costs and benefits on the relative cost-effectiveness and value of a specific exemplar of precision medicine to show how it may be applied to precision medicine more generally.

3.4 Methods

This study uses an adaptation of decision-analytic model-based cost-effectiveness analysis to understand the impact of the value of implementation on precision medicine. This section first describes the two existing (‘traditional’) value of implementation frameworks (VOImp), in both static (VOImp (static)) and dynamic (VOImp (dynamic)) forms, that each assume constant marginal costs and benefits (167). The second part of this section describes how the existing VOImp (static) and VOImp (dynamic) can be modified to allow for varying marginal costs and benefits to take account of the potential impact of capacity constraints on the relative cost-effectiveness of precision medicine. The third section describes how the two modified VOImp frameworks (static and dynamic) were applied to a case study in precision medicine to demonstrate the potential impact of allowing for changes in the marginal costs and benefits when assessing the relative cost-effectiveness of an example test-treat strategy.

3.4.1 Static Value of Implementation with Constant Marginal Costs and Benefits
The value of implementation framework in the static form (VOImp (static)) values the improved net benefit which arises from treating an expanding number of patients following an implementation strategy in a single time-period. The first step in estimating the VOImp (static) involves calculating the net monetary benefit (NMB) of the intervention (see equation 3.1):

\[ NMB = k \cdot \Delta H - \Delta C \]  

(Equation 3.1)

Equation 3.1 shows how a monetary value per additional patient treated is calculated by multiplying the incremental health gained by treating patients with the new interventions compared to a comparator (\(\Delta H\)) by a defined a threshold value (\(k\)) specified for accruing those gains. The threshold represents a cut-off value for the relative cost-effectiveness of the intervention compared with current practice, for example £20,000 per QALY (quality-adjusted life year) gained (68). Any measure of health can be used, but most published value of implementation studies have used the QALY as the measure of health. The incremental cost of treating patients with the intervention rather than the comparator (\(\Delta C\)) is then subtracted from the health gain represented as a monetary value. Note that in the VOImp (static) framework the calculated NMB will be the same for each additional patient, regardless of the level of implementation.

The current value of implementation represents the value of the precision medicine to society in terms of the total incremental net benefit it provides given the proportion of patients receiving the intervention at the current time. This (see equation 3.2) is calculated by multiplying the value of NMB by the total number of patients (\(n\)) in the relevant population and the proportion of those patients receiving the intervention at the current implementation level (\(p\)):

\[ \text{Current Value of Implementation} = n \cdot p \cdot NMB \]  

(Equation 3.2)

The monetary value of increasing to full implementation (see equation 3.3) is then calculated by subtracting the number of patient population currently receiving the intervention (\(1 - p\)) from the NMB multiplied by the total patient population:

\[ \text{Value of Perfect Implementation} = n \cdot (1 - p) \cdot NMB \]  

(Equation 3.3)
The resulting value represents the potential VOImp (static) associated with increasing implementation using a specific strategy. In reality, while it is possible for an implementation strategy to result in full implementation of the intervention, it is likely that the implementation level take some intermediate value. The actual value of implementation represents the additional incremental net benefit that will be provided to society after using an implementation strategy to improve the use of a precision medicine that is cost-effective. To calculate the value of actual implementation (equation 3.4), the degree of uptake of as a result of the implementation strategy ($\sigma$) replaces the assumption of 100% implementation in equation 3.3:

\[
\text{Value of Actual Implementation} = n. (\sigma - p). NMB
\]  
\text{(Equation 3.4)}

The next stage acknowledges the need to take account of opportunity cost by incorporating the resource use and cost for the implementation strategy and noting these will not be available for funding other healthcare services. Equation 3.5 shows how the incremental net benefit of implementation is calculated by subtracting the cost of the implementation strategy ($I$) from the actual value of implementation (equation 3.4):

\[
\text{INB of Implementation} = n. (\sigma - p). NMB - I
\]  
\text{(Equation 3.5)}

### 3.4.2 Dynamic Value of Implementation with Constant Marginal Costs and Benefits

The dynamic current value of implementation (VOImp (dynamic)) outlined by Walker and colleagues takes the form (see equation 3.6):

\[
\text{Current Value of Implementation} = \sum_{t=1}^{T} \frac{n_t \cdot p_t \cdot NMB}{(1 + r)^{t-1}}
\]  
\text{(Equation 3.6)}

The VOImp (dynamic) takes account of the value of implementation in different specified periods of time periods ($t$) measured in annual increments. The time subscripts for the patient population ($n_t$) and proportion of patients currently receiving the intervention ($p_t$) allow for the size of the relevant patient population to change over time and for the provision of the intervention to
naturally change by a process known as diffusion (172). The value gained from the intervention in future years also takes account of the future value of benefits using a discount rate (denoted by \( r \)). The value of perfect implementation (equation 3.7), actual implementation (equation 3.8) and the incremental net benefit of the intervention (equation 3.9), representing the VOImp (dynamic) are then calculated by making the simple substitutions used in the static framework and adding time subscript for \( \sigma \):

\[
\text{Value of Perfect Implementation} = \sum_{t=1}^{T} \frac{n_t \cdot (1 - p_t) \cdot NMB}{(1 + r)^{t-1}} \quad \text{(Equation 3.7)}
\]

\[
\text{Value of Actual Implementation} = \sum_{t=1}^{T} \frac{n_t \cdot (\sigma_t - p_t) \cdot NMB}{(1 + r)^{t-1}} \quad \text{(Equation 3.8)}
\]

\[
\text{INB of Implementation} = \sum_{t=1}^{T} \frac{n_t \cdot (\sigma_t - p_t) \cdot NMB - I_t}{(1 + r)^{t-1}} \quad \text{(Equation 3.9)}
\]

### 3.4.3 Static Value of Implementation for Interventions with Varying Marginal Costs and Benefits

Equations 3.4 and 3.9 show how to calculate the VOImp (static) and VOImp (dynamic), respectively, with the assumption of constant marginal costs and benefits. Marginal costs and benefits are defined as those accrued from treating an additional patient. Incremental costs and benefits are defined as the difference in costs and benefits accrued from using the new intervention compared with a comparator (current practice). The marginal incremental costs and benefits are therefore defined as the additional costs and benefits experienced from treating one more patient with a new intervention rather than current practice. The assumption of constant marginal costs and benefits may provide misleading conclusions about the impact on relative cost-effectiveness of the new intervention if the marginal incremental costs and benefits vary at different levels of implementation. If the marginal incremental costs (\( \Delta C_p \)) and benefits (\( \Delta H_p \)) vary, this will result in a NMB which depends on the level of implementation (equation 3.10):

\[
NMB_p = k \cdot \Delta H_p - \Delta C_p \quad \text{(Equation 3.10)}
\]
The current value of implementation in a static framework with varying costs and benefits (see equation 3.11) is similar to that when using a static framework with a constant NMB (see equation 2):

$$\text{Current Value of Implementation} = n \cdot p \cdot NMB_p \quad \text{(Equation 3.11)}$$

Differences between the constant and variable NMB approaches begin to appear in the equation for calculating the perfect value of implementation (equation 3.12). Since the NMB achieved at perfect implementation ($\sigma = 1$) and current implementation ($p$) will be different, these cannot be collapsed in the way done to produce equation 3.3:

$$\text{Value of Perfect Implementation} = n\left(\sigma \cdot NMB_{\sigma=1} - p \cdot NMB_p\right) \quad \text{(Equation 3.12)}$$

Similarly, separate NMB figures are required when calculating the value of actual implementation that takes change in marginal costs and benefits into account (equation 3.13):

$$\text{Value of Actual Implementation} = n\left(\sigma \cdot NMB_{\sigma} - p \cdot NMB_p\right) \quad \text{(Equation 3.13)}$$

The value of an implementation strategy is now dependent on the initial and final levels of implementation rather than simply being the constant net monetary benefit multiplied by the number of additional patients receiving the intervention. The VOImp (static) features a one-time implementation investment and so a single cost of the strategy ($I$) can be subtracted to find the incremental net benefits of implementation (equation 3.14):

$$\text{INB of Implementation} = n\left(\sigma \cdot NMB_{\sigma} - p \cdot NMB_p\right) - I \quad \text{(Equation 3.14)}$$
3.4.4 Dynamic Value of Implementation with Varying Marginal Costs and Benefits

Varying marginal costs and benefits can also be incorporated into VOImp (dynamic) framework by allowing for the implementation specific levels of NMB (see equation 3.10). The current value of implementation with varying marginal costs and benefits therefore becomes (equation 3.15):

\[
\text{Current Value of Implementation} = \sum_{t=1}^{T} \frac{n_t p_t NMB_p}{(1 + r)^{t-1}} \quad \text{(Equation 3.15)}
\]

It is necessary to take into account that implementation can naturally develop through the process of diffusion. The VOImp (dynamic) will be dependent on the implementation level and the expected relative cost-effectiveness, that represents the added net benefit of the proposed intervention compared with current practice. Therefore, the VOImp (dynamic) can potentially change over time without a specific strategy to change implementation. The value of perfect implementation (equation 3.16) and value of actual implementation (equation 3.17) can be also be formulated by adding implementation level specific NMB:

\[
\text{Value of Perfect Implementation} = \sum_{t=1}^{T} \frac{n_t (NMB_{\sigma=1} - p_t NMB_p)}{(1 + r)^{t-1}} \quad \text{(Equation 3.16)}
\]

\[
\text{Value of Actual Implementation} = \sum_{t=1}^{T} \frac{n_t (\sigma_t NMB_{\sigma} - p_t NMB_p)}{(1 + r)^{t-1}} \quad \text{(Equation 3.17)}
\]

The time (representing a specific time period, for example, year) subscripts on the cost of the implementation strategy \(i\) and it’s the level of implementation following an implementation strategy \(\sigma\) already allow for these factors to be non-linear over time and this may provide an additional source of non-linearity in relative cost-effectiveness representing the added value of the intervention in different time periods (see equation 3.18):

\[
\text{INB of Implementation} = \sum_{t=1}^{T} \frac{n_t (\sigma_t NMB_{\sigma} - p_t NMB_p) - i_t}{(1 + r)^{t-1}} \quad \text{(Equation 3.18)}
\]
3.5 Case Study: 70-Gene Recurrence Score Test for Breast Cancer

This section describes a case study to show the impact of considering the value of implementation in precision medicine while taking account of varying marginal costs and benefits. This example builds on the work of Retèl and colleagues that used a published decision-analytic (Markov) model-based CEA [35] of the 70-gene recurrence score test (Mammaprint ®) to guide treatment selection of early chemotherapy for women at risk of breast cancer recurrence. Mammaprint is a test that provides a score indicating the risk of breast cancer recurrence. This score can be used to inform treatment options which can be stratified based on this score. Women at high risk of recurrence can receive adjuvant chemotherapy while those at a low estimated risk can be monitored without experiencing the potential side-effects of chemotherapy.

Retèl and colleagues, aimed to anticipate the barriers and facilitators to introduce MammaPrint (18) into clinical practice and developed a range of scenarios which might occur during implementation. Retèl and colleagues used the Delphi method to bring together stakeholders in the potential implementation of the Mammaprint test in order to understand the degree of consensus about the potential barriers and facilitators to its introduction and to determine how likely it would be that each barrier or facilitator would occur. An initial list of ten barriers and facilitators were produced and these were then condensed into three critical barriers which were included in the Markov model: technical failure; non-compliance with discordant test results (results which do not align with the clinician’s perceptions of the patient’s risk); and uptake by clinicians. The most significant barrier proved to be the uptake of the intervention (see Table 3.1), which produced significantly different incremental cost-effectiveness ratios at different levels of use of test results by clinicians in decision making. Mammaprint was observed to not meet accepted thresholds of relative cost-effectiveness at the lowest level of implementation (3%) but became more cost-effective with increasing levels of implementation due to rising marginal incremental benefits and falling marginal incremental costs.

The observed varying marginal incremental costs and benefits of Mammaprint arose due to the presence and gradual removal of the barriers to implementation. For example, in the event of technical failure it was assumed that 10% of the cost of the Mammaprint was spent with no concurrent benefit gained. Similarly, when test results were returned but the clinician did not use the result to change practice, a cost was incurred but no benefit gained. These factors result in a higher initial cost per patient of the test when Mammaprint was implemented due to technical
failure and a lower benefit per patient as not all of the tests were used to change clinical practice. As implementation improves over time and these barriers are removed, the marginal incremental costs reduce and the marginal incremental benefits increase, thereby improving the relative cost-effectiveness of Mammaprint.

Table 3.1: Incremental cost-effectiveness ratios of Mammaprint at different implementation levels

<table>
<thead>
<tr>
<th>Assumed proportion of use of test results by clinicians</th>
<th>Marginal Incremental Benefits (QALYs)</th>
<th>Marginal Incremental Cost (€)</th>
<th>Marginal ICER (compared with current practice)</th>
</tr>
</thead>
<tbody>
<tr>
<td>3%</td>
<td>0.001</td>
<td>1,940</td>
<td>€1.9 million per QALY gained</td>
</tr>
<tr>
<td>50%</td>
<td>0.0728</td>
<td>1,630</td>
<td>€22,388 per QALY gained</td>
</tr>
<tr>
<td>92%</td>
<td>0.1492</td>
<td>1,171</td>
<td>€7,853 per QALY gained</td>
</tr>
</tbody>
</table>


3.5.1 A Static VOImp Analysis of an Intervention with Variable Marginal Costs and Benefits

The selected case study used the results of a published state transition Markov model-based cost-effectiveness analysis in which the estimated variation in the marginal costs and benefits were driven by including specific barriers and facilitators to the introduction of Mammaprint into clinical practice (136). The process, driven by the structure and assumptions of the decision-analytic model, of how these barriers and facilitators generate varying marginal costs and benefits is likely to be complex. Ideally, access to the decision-analytic model is required to predict the marginal costs and benefits of precision medicine at different levels of implementation. A working version of the state transition Markov model produced by Retèl et al was not available in the public domain. Therefore, to provide an illustrative example for this study in the absence of a decision-analytic model, the outputs of the analysis produced by Retèl were used. Using these outputs, a simple meta-model was created using ordinary least squares regressions to predict the marginal costs and benefits of the intervention at different levels of implementation of Mammaprint. The three data points for the marginal incremental costs and benefits shown in table 3.1 were used to estimate a linear function (OLS regression) that approximates the benefits and costs as a function of the implementation level ($p$) (equations 3.19 and 3.20). This approximation represents a simple
meta-model of the changes in the marginal incremental costs and benefits which are driven by changes in the more complex underlying Markov model:

\[ \Delta H_p = 0.1662p - 0.006 \]  
\[ \Delta C_p = 1996 - 860p \]  

(Equation 3.19)  

(Equation 3.20)

Equation 3.19 had an \( R^2 \) value of 0.998 and equation 3.20 had an \( R^2 \) value of 0.980. These two \( R^2 \) values indicate the estimated OLS regression was a good approximation of the relationship between the implementation level and the marginal incremental costs and benefits. These equations suggest that as implementation of Mammaprint increases, the marginal health gains will rise and the marginal costs will fall. More generally, a decision-analytic model could be run to generate the estimates of marginal health gains and costs. Using equation 3.20, the variable NMB can then be predicted as a function of the implementation level \( (p) \) by substituting equations 3.19 and 3.20, and a threshold value \( (k) \) into equation 3.10. In this example a threshold value of €20,000 per QALY was used, resulting in a formula for the variable NMB (equation 3.21):

\[ \text{NMB}_p = 4184p - 2116 \]  

(Equation 3.21)

The positive coefficient before the implementation variable ‘\( p \)’ suggests that there are increasing returns to scale associated with Mammaprint. In other words, as the extent of implementation of Mammaprint in the population rises, the net monetary benefit of an additional patient receiving the test-treat strategy also increases. This static VOImp analysis has made the simplifying assumption that that changes in marginal costs and benefits would be driven by the implementation level affecting cost by economies of scale and consequences by learning effects. The analysis also assumed that each patient would receive the same incremental benefit from Mammaprint at the same incremental cost. In practice the incremental marginal costs and benefits of Mammaprint will depend on a range of factors other than the degree of uptake into the population.

The point at which Mammaprint becomes cost-effective, in this scenario, can be calculated by finding the cut-off point of \( p \) when NMB is set equal to zero. Estimating this cut-off point suggests that an implementation level of over 51% is required for Mammaprint to constitute a cost-effective use of resources at a threshold of €20,000 per QALY gained.
The marginal costs and benefits at the highest implementation level presented in Retèl et al (2012) (92%) were used to calculate the NMB when marginal costs and benefits were assumed to be constant to be used as the comparator option. This approach was taken because traditional cost-effectiveness analysis provides estimates which represent the long-run cost-effectiveness of the precision medicine (109). These traditional estimates of cost-effectiveness assume that all patients will receive the precision medicine and factors such as short run fixed costs are not considered (173). The costs attributed to the intervention at the period furthest forward in time from start of the decision analysis therefore best represent the (long-run) results which would be produced by a traditional CEA as they will be as close to the optimal use of the intervention as possible.

3.5.2 The Static Value of a Strategy to Improve Implementation of an Intervention with Variable Marginal Costs and Benefits

The VOImp (static) of any implementation strategy will be dependent on the initial and final level of implementation and not just simply on the magnitude of the change in implementation over time. In the static scenario, when using a constant NMB, a 20% point rise in the degree of implementation of Mammaprint has a constant value of €3,626,000 per QALY gained in a population of 10,000 women (table 3.2). In practice, the non-linear nature of the incremental costs and benefits of Mammaprint will result in a change in the relative cost-effectiveness that is dependent on the initial baseline and final level of implementation. Raising implementation from the baseline value of 20% to 40% had a low level of relative cost-effectiveness, implying that the resources direct to the implementation strategy may have been used to better effect in a different area of the health system. In contrast, moving from a baseline level of 40% to 60% had a substantially bigger impact on relative cost-effectiveness, which in absolute terms meant Mammaprint moving from being a cost-ineffective to a cost-effective use of resources at a threshold of €20,000 per QALY gained.
Table 3.2: VOImp (static) accounting for the baseline and final level of implementation of Mammaprint in a population of 10,000 women

<table>
<thead>
<tr>
<th>Description</th>
<th>Formulae</th>
<th>Constant Marginal Costs and Benefits (N=10,000)</th>
<th>Varying Marginal Costs and Benefits (N=10,000)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Baseline implementation: 20%</td>
<td>Baseline implementation: 20%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Final implementation: 40%</td>
<td>Final implementation: 40%</td>
</tr>
<tr>
<td>Current Value of Implementation Before Introducing an Implementation Strategy</td>
<td>$n \cdot p \cdot NMB_p$</td>
<td>€3,626,000</td>
<td>€7,252,000</td>
</tr>
<tr>
<td>Current Value of Implementation After Introducing an Implementation Strategy</td>
<td>$n \cdot \sigma \cdot NMB_\sigma$</td>
<td>€7,252,000</td>
<td>€10,878,000</td>
</tr>
<tr>
<td>Actual Value of Implementation</td>
<td>$n(\sigma \cdot NMB_\sigma - p \cdot NMB_p)$</td>
<td>€3,626,000</td>
<td>€3,626,000</td>
</tr>
</tbody>
</table>

The baseline level of 20% was assumed using the estimated incidence of breast cancer in the Netherlands of approximately 14,000 cases per year, of which, approximately 80% of women could benefit from Mammaprint (174,175). We therefore assumed approximately 10,000 women a year could benefit from the test.
3.5.3 A Dynamic VOImp Analysis of an Intervention with Variable Marginal Costs and Benefits.

The VOImp (dynamic) of Mammaprint is now shown. It is necessary to make some key assumptions. Specifically, it was necessary to assume that the population of eligible patients does not change over time and there is a diffusion rate of 6 percentage points per year in that population. This assumption reflected the time frame for implementation outlined in Retèl et al. (2012) (136). In this example a discount rate of 3% per year was applied to the net health benefit. Figure 3.1, with the actual values shown in table 3.3, illustrates how the current value of implementation of Mammaprint differed substantially depending on whether constant or varying NMB was assumed.

**Figure 3.1: Current VOImp (dynamic) for Mammaprint assuming constant or variable net monetary benefit**

![Graph showing current VOImp for Mammaprint](image)
Table 3.3: Current Value of Implementation (dynamic) for Mammaprint assuming constant or variable net monetary benefit

<table>
<thead>
<tr>
<th>Year</th>
<th>Constant Net Monetary Benefit</th>
<th>Varying Net Monetary Benefit</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>€ 1,584,175</td>
<td>-€ 1,519,899</td>
</tr>
<tr>
<td>2</td>
<td>€ 2,563,390</td>
<td>-€ 2,104,440</td>
</tr>
<tr>
<td>3</td>
<td>€ 3,484,219</td>
<td>-€ 2,377,955</td>
</tr>
<tr>
<td>4</td>
<td>€ 4,349,233</td>
<td>-€ 2,366,098</td>
</tr>
<tr>
<td>5</td>
<td>€ 5,160,902</td>
<td>-€ 2,093,055</td>
</tr>
<tr>
<td>6</td>
<td>€ 5,921,600</td>
<td>-€ 1,581,619</td>
</tr>
<tr>
<td>7</td>
<td>€ 6,633,607</td>
<td>-€ 853,258</td>
</tr>
<tr>
<td>8</td>
<td>€ 7,299,115</td>
<td>€ 71,824</td>
</tr>
<tr>
<td>9</td>
<td>€ 7,920,227</td>
<td>€ 1,174,623</td>
</tr>
<tr>
<td>10</td>
<td>€ 8,498,966</td>
<td>€ 2,437,277</td>
</tr>
<tr>
<td>11</td>
<td>€ 9,037,273</td>
<td>€ 3,843,010</td>
</tr>
<tr>
<td>12</td>
<td>€ 9,537,013</td>
<td>€ 5,376,077</td>
</tr>
<tr>
<td>13</td>
<td>€ 9,999,975</td>
<td>€ 7,021,714</td>
</tr>
<tr>
<td>14</td>
<td>€ 10,427,877</td>
<td>€ 8,766,089</td>
</tr>
<tr>
<td>15</td>
<td>€ 10,822,370</td>
<td>€ 10,596,252</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>€ 103,239,942</strong></td>
<td><strong>€ 26,390,541</strong></td>
</tr>
</tbody>
</table>

*a assuming a population size of 10,000 patients

Under an assumption of constant NMB, Mammaprint gradually diffuses into practice and provides added value for the health system at a constantly increasing rate, albeit with diminishing returns due to impact of the discount rate. Taking account of the potential for variable NMB it can be observed that, in the initial years following implementation, Mammaprint was not a cost-effective use of resources. Over the first seven years, the total losses of Mammaprint accrue to a total monetary loss of €12,896,324. The resources allocated to Mammaprint over the first seven years could have been used to gain approximately 645 QALYs by funding other interventions in the health system. Mammaprint does become cost-effective from year 9 and produces a marginally increasing net benefit in each year but the total value accrued over the nine years observed is still only 26% of that predicted when using a constant NMB.
3.5.4 The Dynamic Value of a Strategy to Improve Implementation of an Intervention with Variable Marginal Costs and Benefits

Now consider a hypothetical implementation strategy which raises implementation by 3 percentage points in the first year (from 9% to 12%) which is sustained for the following 14 years. Again a discount rate of 3% per year was applied to the net health benefit. The value, in terms of the total incremental net benefit resulting from the implementation strategy, is shown graphically in figure 3.2 and numerically in table 3.4.

Figure 3.2: Actual value of an implementation strategy which improves uptake by 3 percentage points
Table 3.4: Actual Value of Implementation (dynamic) for Mammaprint assuming constant or variable net monetary benefit which improves uptake by 3 percentage points

<table>
<thead>
<tr>
<th>Year</th>
<th>Implementation Level Without Implementation Strategy</th>
<th>Implementation Level with Implementation Strategy</th>
<th>Constant Net Monetary Benefit&lt;sup&gt;a&lt;/sup&gt;</th>
<th>Varying Net Monetary Benefit&lt;sup&gt;a&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.09</td>
<td>0.12</td>
<td>€ 543,900</td>
<td>-€ 371,208</td>
</tr>
<tr>
<td>2</td>
<td>0.15</td>
<td>0.18</td>
<td>€ 528,058</td>
<td>-€ 214,159</td>
</tr>
<tr>
<td>3</td>
<td>0.21</td>
<td>0.24</td>
<td>€ 512,678</td>
<td>-€ 65,944</td>
</tr>
<tr>
<td>4</td>
<td>0.27</td>
<td>0.30</td>
<td>€ 497,746</td>
<td>€ 73,819</td>
</tr>
<tr>
<td>5</td>
<td>0.33</td>
<td>0.36</td>
<td>€ 483,248</td>
<td>€ 205,496</td>
</tr>
<tr>
<td>6</td>
<td>0.39</td>
<td>0.42</td>
<td>€ 469,173</td>
<td>€ 329,441</td>
</tr>
<tr>
<td>7</td>
<td>0.45</td>
<td>0.48</td>
<td>€ 455,508</td>
<td>€ 445,991</td>
</tr>
<tr>
<td>8</td>
<td>0.51</td>
<td>0.54</td>
<td>€ 442,240</td>
<td>€ 555,472</td>
</tr>
<tr>
<td>9</td>
<td>0.57</td>
<td>0.60</td>
<td>€ 429,360</td>
<td>€ 658,197</td>
</tr>
<tr>
<td>10</td>
<td>0.63</td>
<td>0.66</td>
<td>€ 416,854</td>
<td>€ 754,467</td>
</tr>
<tr>
<td>11</td>
<td>0.69</td>
<td>0.72</td>
<td>€ 404,713</td>
<td>€ 844,570</td>
</tr>
<tr>
<td>12</td>
<td>0.75</td>
<td>0.78</td>
<td>€ 392,925</td>
<td>€ 928,785</td>
</tr>
<tr>
<td>13</td>
<td>0.81</td>
<td>0.84</td>
<td>€ 381,481</td>
<td>€ 1,007,378</td>
</tr>
<tr>
<td>14</td>
<td>0.87</td>
<td>0.90</td>
<td>€ 370,369</td>
<td>€ 1,080,604</td>
</tr>
<tr>
<td>15</td>
<td>0.93</td>
<td>0.96</td>
<td>€ 359,582</td>
<td>€ 1,148,711</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td></td>
<td></td>
<td><strong>€ 6,687,834</strong></td>
<td><strong>€ 7,381,619</strong></td>
</tr>
</tbody>
</table>

<sup>a</sup>assuming a population size of 10,000 patients

The value of an implementation strategy which increases uptake by 3 percentage points in each time period is broadly constant in each year when NMB is assumed to be constant. This finding was because the same amount of additional health and cost is gained from treating an additional 3% of patients in each year regardless of the number of patients already receiving the Mammaprint intervention. Under a variable NMB the value of such an implementation strategy is negative in early periods but will increase over time. This finding was because at low levels of implementation the intervention is not cost-effective and so the additional patients receiving the intervention will mean increases in the societal loss arising from implementation. Over time the additional
implementation eventually decreases the marginal costs and increases the marginal benefits for patients meaning that there are increasing returns in later time periods. The overall result is that the implementation strategy is worth nearly €700,000 more when accounting for variable NMB despite the fact that it results in a societal loss soon after its launch.

3.6 Discussion

This study has illustrated a general approach to show how existing value of implementation methods can be adapted to account for varying marginal incremental costs and benefits in a specific case study focussing on precision medicine. The study has adapted existing value of implementation methods, using the outputs from a decision-analytic model-based analysis, to understand the potential impact on the cost-effectiveness of strategies to improve the implementation of precision medicine. This approach differs from traditional approaches, that present the results from a decision-analytic model as incremental cost-effectiveness ratios on a cost-effectiveness plane showing the threshold for cost-effectiveness (65), by taking account of the total number of patients taking up precision medicine as a proportion of the number of patients eligible for treatment (65). Taking account of the level of implementation allows the total societal net benefit of precision medicine to be determined. Calculating the ICER only allows a yes or no decision to be made regarding acceptance of precision medicine dependent on a stated threshold (68). Precision medicine can appear to be cost-effective but not maximise total societal net benefit if it is only implemented in a proportion of the total eligible patient population.

In this study we have built on a published decision-analytic model-based cost-effectiveness analysis of Mammaprint and shown that assuming constant marginal net costs and benefits may produce misleading results when assessing the relative cost-effectiveness of implementation strategies. The need to take account of the impact of implementation on relative cost-effectiveness is important to any new healthcare intervention. However, the exemplar of precision medicine has specific and known multiple barriers to timely and effective implementation (91,101). Varying marginal costs and benefits may be a particularly prevalent issue when introducing exemplars of precision medicine. Such interventions often involve some kind of test or stratifying mechanism to identify which patients may benefit or be harmed by a treatment or to predict disease course. Previous studies have identified a lack of testing capacity as a barrier to precision medicine and other issues such as the need to develop logistical mechanisms to deliver tests in a timely manner may also impede implementation (89,90,101,102). As such it is possible that the significant upfront
investments in capacity would mean that it would not be cost-effective to use the precision medicine to treat a small eligible patient population. Another barrier may be a lack of training or guidelines in the use of precision medicine and as such there may also be learning effects where the outcomes of testing and treatment improve as test providers and clinicians gain experience in providing the interventions (90,91).

The relative cost-effectiveness and net marginal value of precision medicine will depend on the level of implementation. Without using implementation strategies, it is possible that losses will accrue over multiple time periods before implementation reaches a cost-effective level. Information about the required level of implementation would not be provided by a conventional CEA, which would have judged the precision medicine to be universally cost-effective based on the average ICER calculated for treating all of the patients in the long run.

It is necessary, but not sufficient, to use estimates of the long-run cost-effectiveness to understand whether an intervention is an efficient use of healthcare resources. In some instances, the decision problem is more appropriately expanded to consider the implementation of interventions that appear to be cost-effective. This is likely to be of particular relevance in precision medicine which involves the implementation of two elements, a ‘test’ and a ‘treatment’ into existing pathways of care. In such instances, estimates of cost-effectiveness should allow for varying marginal costs and benefits and differential cost-effectiveness at different levels of implementation of precision medicine. These approaches could use static models with costs and benefits that are described as a function of implementation levels and producing estimates of cost-effectiveness at different levels or dynamic, multi-cohort models to capture the impact of changing marginal costs and benefits over time (161).

In applied studies using the methods outlined in this paper, the incremental costs and consequences of a specific example of precision medicine will be estimated using a decision-analytic model. Using the framework of a decision-analytic model will allow the analyst to vary the marginal costs and benefits depending on pre-specified factors such as capacity constraints, economies of scale, learning effects, and patient characteristics. The computational effort, in terms of technical skills and computer processing power, in adapting existing decision-analytic model-based cost-effectiveness analysis is minimal. The approach involves making modest adaptations to the structure of existing decision-analytic models and presenting the outputs in a particular way. However, this approach does require an understanding of potential barriers and facilitators to
implementation, which may require mixed-methods approaches such as the Delphi method or qualitative methods such as semi-structured interviews (176,177).

The method proposed in this paper will be relevant to a number of applications and specified decision problems. Clear criteria will be needed to define the scope of the relevant decision problem that need to take account of the value of implementation. Such criteria will be needed to identify potential interventions which are likely to exhibit varying marginal incremental costs and consequences. For example, vaccines exhibit positive externalities that arise because the vaccine protects not only the individual against infection but also other people within a population because the spread of disease can be halted. The magnitude of the additional benefits accrued will reduce as more people are vaccinated and therefore do not benefit from the externality. Another potentially relevant example is in the implementation of a new surgical technique. As surgeons gain experience in the surgical technique the length of time required to perform the operation, and hence costs, may reduce in addition to improvements in patient outcome. These two examples illustrate some potential situations in which marginal incremental costs and benefits may arise indicating that the method proposed in this study may be valuable.

Other causes of variable marginal costs and benefits in an intervention include: economies or diseconomies of scale may impact the marginal cost per patient as implementation increases (77); economies of scope may arise as resources can be shared over multiple interventions (77); learning curves may mean that clinicians’ experience and skill in using an intervention increases as they treat more patients and therefore patient health outcomes increase (79); prioritisation of sicker patients may lead to decreasing marginal health benefits (80).

The issue of variable marginal costs and benefits has previously been explored by Lord et al., (2006) with regards to the implications of non-linear cost-effectiveness frontiers for funding decisions (80). For a convex frontier, for example for an intervention where patients are prioritised by likely benefit leading to decreasing returns to scale, the optimal decision may actually be to fund a portion of the new intervention while keeping some patients on the old intervention. These authors found that when phenomena such as decreasing marginal costs caused the frontier to be concave then the decision to either fund the intervention for all patients or to fund it for none would always be the optimum in terms of net benefit. They concluded that “partial implementation will not be cost-effective, and hence estimation of the expansion path (the movement from the old to the new intervention) will not be productive” (80). However, it is possible that an imperfect
implementation of an intervention which is cost-effective at population level will be cost-effective when compared with only funding the existing intervention. Similarly, some levels of implementation of the new intervention will not be cost-effective. Whether the implementation level of the new intervention is cost-effective or not will depend on the point at which the cost-effectiveness frontier crosses the threshold of relative cost-effectiveness (for example £20,000 per QALY gained) for the health system. Therefore, while partial implementation will never be the most cost-effective combination of the two interventions in concave frontiers, in reality the level of implementation is critical to the cost-effectiveness of the intervention.

One previous study by Grimm and colleagues allowed for varying marginal costs in assessing the cost-effectiveness of implementation of a preterm birth screening test but not in a value of implementation framework (178). Grimm et al., (2016) modelled the impact of falling future prices on the cost-effectiveness of the screening test. In their case study, the cost of the screening test fell to 90% of the previous value every time the number of patients receiving it doubled. The incremental cost-effectiveness ratio (ICER) reduced by up to 46% with increasing implementation. The authors suggested that to account for this varying marginal cost in decision making, a dynamic ICER could be produced reflecting the average ICER across all time periods but did not go on to provide this analysis.

**Limitations**

This study has used an example of precision medicine where changes in the estimated marginal costs and benefits were modelled, in the absence of access for a fully executable decision-analytic model, as approximately linear functions of changes of implementation level. While this simplified the mathematical solution to the problem, and is useful in demonstrating the consequences of varying marginal costs and benefits, in reality marginal costs and benefits may not be a smooth function of implementation. For example, there may be a number of capacity investments that are required to increase patient access to elements (test and/or treatment) of precision medicine. Each capacity investment has its own cost and effect in terms of increasing implementation. Over the course of achieving full implementation, a current value of implementation curve (for example, figure 3.1) may, in practice, comprise a number of different, linked curves with different slopes. While there may be an apparent overall effect of increasing or decreasing net monetary benefit, a formula to predict such changes by modelling NMB as a function of implementation may not represent a good approximation. In such cases, the marginal costs and benefits at each
implementation level of interest may have to be estimated directly based on input data or a more complex underlying model such as a discrete event simulation.

A key assumption made in the example presented in this paper is that the required parameters were known with absolute certainty. In addition the case study took the perspective of determining the value of implementing an intervention after it has been approved as cost-effective based on estimates of the cost-effectiveness of Mammaprint at full implementation. The aligns with current approaches to HTA which implicitly take a “long run” perspective while failing to account for differential cost-effectiveness in the short term (109,111). Given that the decision to adopt has been taken based on certain evidence, the counterfactual used in this example is continued use of the comparator which produces no incremental net benefit.

In practice, accounting for uncertainty is critical when determining the value of implementing interventions and whether it is necessary to obtain more information about a specific parameter through further research. It is difficult to quantify the value of an implementation strategy when there is a probability that the intervention itself is not cost-effective. The presence of varying marginal costs and benefits will have significant and complex consequences for such applications of VOImp and VOI methods by introducing an additional layer of uncertainty. In addition to estimating the distribution of costs and benefits for the intervention and comparator, the analyst will also have to forecast how these costs and benefits will change depending on implementation levels. In the example presented in this paper, the levels of implementation for which the intervention would not be cost-effective due to high marginal costs and low marginal benefits would also be associated with a degree of uncertainty. Implementation strategies will only be shown to be of added value if there is a sufficiently large probability that the strategy will result in a distribution of implementation levels at which introducing the precision medicine becomes cost-effective at a pre-defined threshold.

In the presence of uncertainty, value of information methods can be used alongside VOImp to determine whether it is better to invest in implementation or better evidence as to the values of the parameters in the model (168,170,179). Such approaches determine the optimal combination of research and implementation of an intervention compared to the counterfactual of implementing the comparator with the highest net benefit given current levels of information (169).
The presence of varying marginal costs and benefits will also have an impact on the expected value of further research for precision medicine. For example, there may be significant value in identifying the levels of implementation at which precision medicine would not be cost-effective by providing better estimates of the marginal costs and benefits. This would minimize the risk of a resulting loss to society by introducing a strategy to deliver precision medicine that was not cost-effective or not cost-effective at the observed degree of uptake in clinical practice.

This study used an example of precision medicine with marginally decreasing costs and increasing benefits, with the overall effect of producing increasing net monetary benefit with increasing implementation. However, it is possible that different patterns of non-linearity are exhibited in marginal costs and benefits. For example, a strategy to deliver precision medicine with marginally increasing costs and/or decreasing benefits will have an opposite pattern of total net benefit with implementation. At lower levels of implementation, precision medicine will be more cost-effective, potentially becoming cost-ineffective at higher levels due to the decreasing returns to net monetary benefit. For example, prescription of oral pre-exposure prophylaxis (PrEP) in individuals at risk of infection with human immunodeficiency virus (HIV) provides a benefit to the patient in that the risk of infection is dramatically reduced but may also provide a positive externality in preventing that individual from passing on the virus to others (149). However, the use of PrEP may be more beneficial in those patients whose behaviours put them at greater risk of infection and as such the cost-effectiveness of the intervention may decrease with the reducing risk status of the patients treated. As such, a strategy to stratify the population by risk of infection may be useful to ensure the intervention is offered in a cost-effective manner.

For some examples of precision medicine, with decreasing returns to scale, there may be concerns about the equity of only providing the intervention to a smaller number of individuals in order for provision of precision medicine to remain cost-effective. While the equity-efficiency trade-off is an extensive existing research area (180–182), the value of implementation approach outlined in this paper may help to quantify the total loss in net benefit which would be incurred from increasing implementation from the maximum cost-effective level to full implementation.

It is also conceivable that an intervention could exhibit both increasing and decreasing returns to net monetary benefit at different levels of implementation of precision medicine. For example, a strategy to deliver precision medicine involving mutation testing to guide a cancer treatment may have high initial marginal costs due to the need to invest in testing equipment and training. As
these costs are divided across increasing numbers of patients, their impact decreases, making the intervention more cost-effective and creating a positive total net benefit. However, higher marginal costs and slower turnaround times may be faced in sending samples from rural hospitals to laboratories in cities. This may lead to decreasing marginal returns when treating these patients and achieving full implementation. In such circumstances, it is possible that there is both a minimum and maximum level of efficient implementation for the intervention.

It is also important to note that the pattern of non-linearity in the marginal costs and benefits of any existing interventions which are to be disinvested from to fund precision medicine will also have implications for the efficiency of the health service. If an existing intervention to be de-funded exhibits increasing returns to scale, then the initial disinvestment will cause a significant loss in health for minimal cost savings. An intervention with decreasing returns could be de-funded with larger savings gains for smaller health losses.

3.7 Conclusion

This paper and selected case study focussing on precision medicine have shown the importance of going beyond an evaluation of the long run cost-effectiveness of an intervention and the limiting assumption of a constant net monetary benefit. Anticipating the pattern of non-linearity in marginal costs and benefits and its impact on the value of implementation is important to ensure that otherwise cost-effective interventions are not implemented in a way which causes a net health loss to society. The use of value of implementation analysis to evaluate implementation improving strategies is also likely to produce significantly biased results if varying marginal costs and marginal benefits are not incorporated.

3.8 Summary of Chapter 3

Chapter 3 has shown that the varying marginal costs and benefits of an example of precision medicine can impact its cost-effectiveness at different levels of implementation and failing to account for these variations may lead to biased estimates of the value of investments in improving implementation.

When introducing a new example of precision medicine, a key element will be to be able to identify potential capacity constraints which may impede its implementation and impact on the
cost-effectiveness of the intervention. Chapter 4 presents an example of a qualitative interview study which sought to identify the relevant barriers and capacity constraints to the introduction of existing examples of precision medicine for non-small cell lung cancer.
Chapter 4
Understanding Barriers to the Introduction of Precision Medicine in Non-Small Cell Lung Cancer: A Qualitative Interview Study

4.1 Chapter 4 Summary

The results of the empirical study described in chapter 3 indicated that it is necessary to be able to identify the relevant capacity constraints that will influence the estimated cost-effectiveness of precision medicine. Chapter 4 reports an empirical study that uses qualitative research methods. This empirical study used a specified case study relevant to precision medicine: the introduction of test-treat strategies for the management of non-small cell lung cancer (NSCLC). The protocol for this empirical study was published: Wright S, Daker-White G, Newman W, Payne K. Understanding barriers to the introduction of precision medicines in non-small cell lung cancer: A qualitative interview protocol. Wellcome Open Research. 2018;3:24 (see Appendix 4.1).

The aim of this study was to explore the type and extent of barriers experienced by service providers and service commissioners when introducing precision medicine for NSCLC for relevant individuals within specific patient populations. The motivation for this study was to address the lack of evidence about the barriers which were faced in introducing examples of precision medicine for NSCLC in the UK. The design, analysis and discussion of the implications of the results of this study using qualitative methods are now presented in chapter 4. The methods sections (sections 4.2) that follow have updated the proposed methods described as published in Wright et al (2018).
4.2 Introduction

Recent developments in the treatment of non-small cell lung cancer (NSCLC) have focussed on moving from using chemotherapy-based regimens to medicines which can target specific mutations in a tumour. The first such class of medicines to be recommended by the National Institute for Health and Care Excellence (NICE) for first-line use in NHS England were the EGFR tyrosine kinase inhibitors (TKIs), beginning with gefitinib in 2010 (100). In patients with tumours that showed EGFR mutations, treatment with EGFR TKIs has been shown to extend the length of time a tumour takes to progress to a life-threatening size by two to five months (35,36). Furthermore, EGFR TKIs may offer greater improvements in quality of life and fewer severe side effects than standard chemotherapy (36,183). Following the introduction of gefitinib other targeted therapies have been developed and introduced into practice. For the approximately 5% of patients whose tumours exhibit ALK mutations, crizotinib can be used as a first line treatment (184). Treatment with crizotinib has been shown to improve patient’s progression free survival by over 4 months (185) compared with standard chemotherapy. Treatments targeting a number of other biomarkers including PD-L1 overexpression or ROS1 alterations have also been recommended by NICE and more are currently under evaluation (186,187)

While these examples of precision medicine offer the potential to improve patients’ quality and length of life, there is evidence that their introduction into the NHS has been slower than anticipated. A significant issue raised in the NICE technology appraisal of the medicines targeting EGFR mutations was that EGFR mutation testing was not current practice in the NHS before the medicines were available. In the 2010 appraisal of gefitinib, EGFR testing was not widely available in the NHS, although it was predicted that it could be quickly implemented with significant investment (100). In the 2012 appraisal of erlotinib it was stated during the NICE technology appraisal process that EGFR testing had become best practice. In 2013, the NICE Diagnostic Assessment Programme conducted an evaluation of testing methods for EGFR, which highlighted that there was heterogeneity in how testing was provided (61,162). A report published in 2014 by Cancer Research UK estimated that 48% of patients eligible for a targeted treatment for NSCLC were not receiving tests (99). As a result, 1,429 out of 3,007 patients who could have benefitted from targeted treatments were estimated to be missing out.

Variation was also being experienced in the turnaround time for tests meaning that some patients began treatment before receiving their results (162). These concerns have been repeated in peer-
reviewed published literature and NHS reports (88,188) and a recent survey sponsored by the pharmaceutical company Boehringer-Ingelheim (189).

When NICE approves a treatment, as part of the technology appraisal programme, it must be made available for all members of the eligible patient group within 3 months and paid for by service commissioners (82). When NICE evaluates a diagnostic technology, guidance about its use is produced but there is no legal requirement to fund and provide the test (70). There is evidence to suggest that testing is not immediately available for the entire patient population and may require years to develop (99). Even as testing is made available for all individuals, differences in the quality of testing and the turnaround time may have implications for treatment decisions meaning that the cost-effectiveness of test-treat interventions when used in NHS practice may differ from the cost-effectiveness estimated at the time of appraisal.

Previous research has shown that there have been issues with implementing examples of precision medicine for lung cancer but few have explored why this was the case for these test-treat interventions in particular (90). One pilot study, conducted to inform a larger qualitative interview study, used face-to-face semi-structured interviews which explored how oncologist’s perceptions and work environment affected their use of genomic-targeted medicines in clinical practice in the United States (90). The published protocol for this study presents the results of a qualitative pilot study (90). Approximately a third of the ten oncology fellows interviewed in the pilot study were uncertain about guidelines regarding the use of precision medicine as second or third-line treatments for lung cancer while a third of those interviewed were also uncertain regarding how to order testing (90). Common barriers to performing tests included insufficient tissue samples, the inconvenience of testing and the cost of testing. Facilitators of tests were the ease of testing and deciphering results, as well as patients having health insurance. The cost of treatment was mentioned as a barrier by a smaller number of clinicians.

A number of additional qualitative studies have sought to identify the barriers to precision medicine beyond NSCLC. In a 2013 study based in Canada, which used focus groups to explore the views of physicians about the future role of personalised medicine in health care, eight key relevant concerns about introducing personalised medicine were identified. These eight concerns were: insufficient knowledge; a need for training of physicians; lack of specific guidelines and protocols for using tests; unequal access to testing due to socioeconomic differences; the financial burden of testing on public funds; additional time pressures that precision medicine will put on
clinical practice; need for geneticist support after testing (91). In a qualitative study that used interviews to identify the barriers to introducing a test for the BRCA gene which significantly raises women’s risk of breast cancer, clinicians highlighted potential issues with co-ordinating the receiving of test results with the timing of treatment decisions (102). It was perceived by clinicians that the barriers to introducing BRCA testing could result in “postponement or avoidance of tests, delayed treatment decisions, and proceeding with decisions before test results [were received]”.

Despite the number of studies investigating barriers to the uptake of precision medicine in general, there has been a paucity of research focusing on the delayed implementation of precision medicine for NSCLC. There has been no study examining the barriers to implementing precision medicine for NSCLC in NHS England. This study sought to create a typology of the organisational barriers to the introduction of examples of precision medicine in NSCLC. Organisational barriers are those that arise due to the way in which the health system operates. These organisational barriers are in the control of the health system and so could be removed or reduced by taking actions to improve implementation. This typology was be used to inform the barriers included in a decision-analytic model-based cost-effectiveness analysis of a test-treat intervention that accounts for capacity constraints (see Chapter 6). It may also be useful as a checklist which can be used by decision makers to anticipate potential barriers to the introduction of further examples of precision medicine.

4.2.1 Aims and Objectives

The primary aim of this study was to explore the type and extent of barriers experienced by service providers and service commissioners when introducing precision medicine for NSCLC for relevant individuals within specific patient populations.

This study had four objectives to explore the views of stakeholders in the provision of examples of precision medicine for NSCLC to identify:

- the types of perceived organisational barriers to introduce examples of precision medicine for NSCLC in NHS England;
- the potential impact for NHS patients of the identified different barriers to the provision of licensed test-treat medicines indicated for the treatment of NSCLC;
- how the availability of existing licensed test-treat medicines indicated for the treatment of NSCLC has changed over time;
• a typology of barriers which may apply to introducing precision medicine beyond NSCLC.

4.3 Methods

This study used semi-structured telephone interviews with service providers (clinicians and test providers) and service commissioners to identify the barriers to introducing precision medicine for NSCLC in NHS England. This study was approved by The University of Manchester Proportionate Review Research Ethics Committee (Reference number: 2017-1885-3619; 25/08/2017).

4.3.1 Rationale for Using Qualitative Research Methods

In this study, the aim was to explore the experiences of service providers to identify the type and extent of barriers to introducing examples of precision medicine for NSCLC. This represents a “what” question as outlined by Green and Thorogood (2018). Qualitative methods were therefore, deemed most appropriate to address the study aim. Quantitative research methods, in contrast, seek to collect and analyse typically numerical data to assess “preconceived models, hypotheses, or theories” (190). They are therefore useful for addressing specific types of research questions such as “how much” or “how many” but less effective at addressing the “what”, “why”, or “how” questions which seek to understand the mechanisms for phenomena (191).

Qualitative methods are inductive in that the researcher typically draws on understanding from the collected data rather than using data to test a pre-existing theory. This approach means that qualitative methods they are effective at exploratory research which is aimed at generating new insights. The field of implementation science offers potential qualitative methods to understand the impact of moving complex interventions into health care systems [see also section 4.5]. However the focus of this PhD was to use theory and methods based in the field of health economics which has only recently begun to recognise the potential contribution of qualitative research methods (192). The qualitative study reported in this chapter will seek to identify the barriers to using examples of precision medicine for NSCLC by exploring the experiences of service providers in implementing such interventions in practice. Quantitative methods will be used in other empirical studies reported in this thesis (Chapters 5, 6, and 7) designed to understand “how much” of an impact the identified capacity constraints have on the cost-effectiveness of examples of precision medicine for NSCLC.
A number of methods are available to collect qualitative data. For example, focus groups can be used to bring multiple participants together in person to discuss an issue. An advantage of focus groups can be that bringing multiple types of participant (such as patients and clinicians) together can yield richer information through communication between the participants. For example, the inclusion of multiple participants allows for agreement or dissent on topics of discussion and potential elaboration on reasons for beliefs that would not be as easily identified in one to one interviews (193). However, this reliance on interaction can also represent a disadvantage of focus groups in the potential for some individuals to dominate the discussion in the group.

Interviews offer an alternative method of qualitative data collection. Interviews involve a typically one to one discussion which can be conducted in person or via the telephone. An interview schedule containing a list of questions or topics is developed to guide the interview and this can range from a very structured to an open ended list of questions. Qualitative interviews may offer an advantage in allowing the researcher to address more sensitive topics which participants may not comfortable discussing in a focus group (194,195). Participants who may be less willing to speak up in a group discussion may also be more comfortable discussing a topic in private with a single interviewer. On a more practical level, it has been argued that while interviews may take longer than focus group discussions, they are able to reach data saturation quicker and offer a more cost-effective approach to qualitative data collection (196). However, a disadvantage of interviews may be that the discussion may be more closely reliant on the interview schedule with less scope for new areas of interest to arise than in focus groups where group dynamics occur (197). However, a review of empirical studies which investigated the potential differences in the results of interviews and focus groups found that the two methods provide broadly similar results (198).

In this study, telephone-based semi-structured interviews were used to collect qualitative data due to the focus of this work on capturing a geographically diverse sample to represent heterogeneity in health care provision. Telephone interviews offer similar advantages to face-to-face interviews while allowing more flexibility in arranging the timing of the interview.

4.3.2 Sample Frame

The sampling frame aimed to identify stakeholders with experience of introducing precision medicine for NSCLC. The relevant stakeholders where identified by examining the NICE care pathway for patients with NSCLC (199). The relevant stakeholders were drawn from three groups:
clinicians; test providers, for example pathologists and geneticists; and service commissioners which may include individuals who are members of care commissioning groups or those involved in commissioning at the national level through NHS England. The principle service providers of interest were oncologists and respiratory physicians specialising in lung cancer but also geneticists and pathologists who are key in providing examples of precision medicine for NSCLC. Examples include EGFR, ALK and PD-L1 testing for medicines such as erlotinib, ceritinib and pembrolizumab.

While factors linked to demand for precision medicine by patients, such as uptake of testing or treatment or adherence to medicines treatment may also impede the implementation of precision medicine, the focus of these interviews was to identify potential supply side capacity constraints. This is because these capacity constraints in the supply of testing or treatments are within the control of the health system. For this reason, patients were not interviewed in this study.

Clinicians and test providers with over seven years of NHS experience were targeted as such individuals were more likely to have direct experience of the introduction of EGFR and ALK testing and treatment as they were working in clinical practice. Clinicians and test providers were recruited via the British Thoracic Oncology Group (BTOG) [15] and the Royal College of Pathologists (RCPat) [16]. Details about the study and an invitation to participate were circulated via the BTOG mailing list which currently has 2083 members and the RCPat list that has over 11,000 members.

The targeted service commissioner sample comprised hospital, regional and national level individuals involved with service commissioning and funding decisions. Examples of service commissioners may involve members of care commissioning groups, hospital finance staff and decision makers involved with national organisations, such as NICE. As service commissioners were likely to come from a range of organisations, there was no universal sampling frame available to reach them. Service commissioners were therefore recruited using existing links and collaborations within the supervisory team to identify an initial sample. As in recruitment for the clinician sample, geographical diversity was sought through purposive sampling and service commissioners were required to have been in a relevant position when EGFR and ALK mutation based testing and subsequent treatment were introduced.
4.3.3 Sample Technique

Purposive sampling was used to gain a diverse sample in terms of the setting and geographical location of testing and treatment (200). These characteristics are likely to be important in the context of introducing examples of precision medicine as experiences may vary depending on the size and nature of hospitals. For example, mutation testing services may be more readily available in larger teaching hospitals with established links to laboratories. For smaller, general hospitals there may be a greater logistical challenge in sending samples for testing and receiving results in a timely manner.

4.3.4 Sample Size

There are no defined rules for calculating sample size in qualitative studies (201). In quantitative studies, a sufficient sample size is required to identify statistically significant differences in the variables of interest. However, interviews intend to identify the breadth of experiences, thoughts, or opinions on a given subject. This study therefore started with an approximate sample of 10 clinicians or test providers, and 10 service commissioners but sampling continued iteratively until no new themes arose from the collected data, that were analysed as new data were collected, otherwise known as inductive thematic saturation (202).

4.3.5 Recruitment

Information regarding the study was sent to clinicians and test providers using mailing lists, with contact details of the principal investigator provided for those interested in taking part. The individuals that expressed an interest in taking part were then subsequently sent more detailed information about the study. Service commissioners were directly sent an email including information about the study and the contact details of the principal investigator. Snowball sampling was used for both samples whereby participants were asked if they knew any other individuals who meet the inclusion criteria who may have been interested in taking part in the study (203).

Clinicians and service commissioners who were interested in taking part in the study were asked in the mailing list adverts to email or phone a named individual (SW) to express an interest in taking part. The researcher then emailed the potential participant a participant information sheet [Appendix 4.2]. After receiving an information sheet, potential participants were given at least 24
hours to consider taking part in the study. If they agreed to take part they were asked to complete a written consent form and to return a copy to the researchers by post or email [Appendix 4.3].

4.3.6 Data Collection

Semi-structured interview schedules were created for the two study samples. The interview schedule was piloted with two clinicians before study recruitment began. The interview schedule for service providers (see Appendix 4.4), and service commissioners (see Appendix 4.5) were, informed by the systematic review of previous economic evaluations of precision medicine (including health technology assessments) reported in chapter 2, and consultation with two expert clinical advisors who are lung oncologists. The core questions for each of the two interview schedules were similar, there were slight variations in the way questions are asked depending on the particular role of the interviewee. For example, clinicians were asked primarily about their experience offering treatments to patients while for geneticists and pathologists the focus was on offering testing.

All telephone interviews were conducted by one researcher (SW) at The University of Manchester and were digitally-recorded. The recordings were then transcribed verbatim by a contracted transcription company called 1st Class Secretarial Services (204). Recordings were sent via an encrypted data transfer.

4.3.7 Data Analysis

The objectives of the data analysis were to: create a typology of barriers which may prevent patients’ access to precision treatments for NSCLC; determine which were the most important barriers; and to identify strategies to improve the implementation of precision treatments. The qualitative data were analysed using a framework analysis facilitated by NVivo software (205). Framework analysis is a five stage process involving; familiarisation; identifying a thematic framework; indexing; charting; mapping; and interpretation (206).

In the initial familiarisation stage, one researcher (SW) transcribed read the first three interviews in order to gain an in-depth understanding of the initial themes emerging from the data. The initial key themes identified during the data familiarisation stage, alongside evidence from previous research (see Chapter 2), formed an initial thematic framework against which the selection of data
was sorted and collected (207). As semi-structured interviews were used for this study, many of the themes originated in the questions contained in the interview schedule. As new themes were identified in the data, they were added to the framework. Each transcript was then indexed against these themes, with sections from the text which support different themes annotated for later retrieval. In the context of this study, the identified themes were the range of barriers which occur in providing and accessing examples of precision medicine for NSCLC and views about which barriers were most significant in restricting the provision of precision medicines.

### 4.3.8 Data Storage and Anonymisation

The recording device and digital memory card containing the interview recordings were stored in a locked draw in a secure university office. The recordings were saved onto an encrypted university computer and the files password protected. Recorded interviews were deleted from recording devices after they have been stored on a computer and anonymised and will be destroyed completely at the end of the study. Interview transcripts will be stored for ten years.

Anonymisation was accomplished by removing references to participants’ names as well as any reference to information which could lead to identification of the participant such as the name of their place of work. When referencing data from the transcripts, generalised information regarding the participant is provided to demonstrate their demographics whilst not allowing identification.

### 4.4 Results

The section reports the results from 11 interviews with participants including; five clinicians, three pathologists, two clinical geneticists, and one service commissioner. All of the clinicians were consultant oncologists, the pathologists comprised two consultant histopathologists and a biomedical scientist, the geneticists were clinical scientists, and the service commissioner was a consultant in clinical genetics. Participants were based in a range of hospitals including city-based teaching hospitals, city-based general hospitals serving town and rural communities. The interviews took place between March 2018 and October 2018 and each interview lasted a mean of 23 minutes (range: 10 minutes to 39 minutes). Data saturation was achieved in the clinician sample but it was not possible to recruit sufficient numbers of test providers or service commissioners to achieve data saturation and as such results from this sample can only be viewed as indicative.
A total of 17 barriers to introducing precision medicine for NSCLC were identified. These 17 barriers were grouped into five key broad themes (see table 4.1): the managed entry of precision medicine for NSCLC; the commissioning and reimbursement of precision medicine for NSCLC and specifically the test component of precision medicine; the complexity of the logistics around providing tests; opinions about whether test provision should be localised or centralised; and opinions about future developments, including potential barriers to their introduction, in precision medicine for NSCLC. None of the participants identified any facilitators to the introduction of precision medicine for NSCLC.
Table 4.1: A Typology of Barriers to the Introduction of Examples of Precision Medicine for NSCLC

<table>
<thead>
<tr>
<th>Theme</th>
<th>Barrier</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Managed entry of precision medicines for NSCLC</td>
<td>Delays between the end of trials or early access to medicines programmes and NICE approval</td>
<td>While NICE is appraising whether a targeted treatment should be made available to all patients in the NHS, there is no funding available to make it available to new patients.</td>
</tr>
<tr>
<td>Increasing number of targeted treatments being appraised by NICE</td>
<td></td>
<td>NICE is appraising an increasing number of targeted treatments for patients with NSCLC and this may lead to a slower appraisal process and delays in new patients accessing the medicines</td>
</tr>
<tr>
<td>Funding mechanisms and service commissioning</td>
<td>Withdrawal of pump-priming funding for biomarker testing</td>
<td>If pharmaceutical companies withdraw funding for testing before the NHS can provide enough tests, some patients may not receive treatments targeted to their mutations or may experience delays in receiving treatment</td>
</tr>
<tr>
<td>Geographical inequalities in access to funding for testing</td>
<td></td>
<td>In the absence of a clear route for reimbursement of test costs, smaller or more rural hospitals may find it more</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td></td>
<td><strong>difficult to provide testing than larger hospitals.</strong></td>
<td><strong>No mandate for NHS England to fund testing when a targeted treatment was approved</strong></td>
</tr>
<tr>
<td></td>
<td>Upon approval of a targeted treatment by NICE, the cost of the drug will be reimbursed to hospitals by NHS England. However, previously there was no such provision for the test required to target the treatment. As such hospitals often offered the test at a loss as they didn’t know how the test was reimbursed. There was conflict between different service providers about who should pay for the test.</td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>Lack of awareness of guidelines about test reimbursement</strong></td>
<td><strong>As of 2016, most biomarker tests required for a targeted treatment approved by NICE have a clear reimbursement route. However, there is still a lack of awareness of these guidelines.</strong></td>
</tr>
<tr>
<td></td>
<td><strong>Lack of funding for test validation</strong></td>
<td><strong>There is a lack of funding to validate the locally conducted immunohistochemical testing required for some targeted treatments. This may lead to heterogeneity in the quality and therefore ability to stratify of the biomarker testing in different areas.</strong></td>
</tr>
<tr>
<td>Logistics of Organising Testing to Guide Treatment</td>
<td>Delays occurring in pathology laboratories have knock on effects for other testing</td>
<td>Pathology laboratories are involved at the beginning of the testing pathway in preparing samples and conducting immunohistochemical screening. If there are delays at this stage, this causes delays for subsequent genetic testing and in returning results to patients</td>
</tr>
<tr>
<td>Poor quality of samples sent by pathology</td>
<td>If the quality of samples sent by pathology to genetic laboratories is insufficient then there may be delays as an additional tumour sample will have to be requested and resent. Patients may refuse a second biopsy meaning that testing cannot take place</td>
<td></td>
</tr>
<tr>
<td>Increasing workload for pathology laboratories</td>
<td>Pathology laboratories are required to process samples, provide a number of immunohistochemical tests and send samples to different laboratories for genetic testing. As the number of biomarker tests required increases, pathology laboratories may find it hard to keep up with this demand, resulting in a slower turnaround time for tests</td>
<td></td>
</tr>
<tr>
<td>The quantitative nature of the PD-L1 test</td>
<td>When conducting the immunohistochemical test for PD-L1 overexpression, pathologists must estimate the percentage of the tumour sample showing overexpression. This introduces subjectivity into the result. If pathologists do not analyse many tests then they may be less effective at estimating this percentage.</td>
<td></td>
</tr>
<tr>
<td>Understaffing of pathology laboratories</td>
<td>Even when funding is available for testing, a shortage of staff in pathology laboratories may slow the turnaround times for tests.</td>
<td></td>
</tr>
<tr>
<td>Sequential biomarker testing</td>
<td>Currently many hospitals run the tests in a sequence. This means that there is a delay in waiting for the initial test results to be returned before additional tests are requested. It is also possible for pathology laboratories to run out of tumour sample to be tested for the latter tests requiring additional samples to be taken.</td>
<td></td>
</tr>
<tr>
<td>Tests conducted in different locations</td>
<td>For hospitals which do not conduct testing in-house, the samples may have to be sent to different hospitals for different</td>
<td></td>
</tr>
</tbody>
</table>
tests. This when combined with the sequential testing of samples may further add delays to receiving results.

<table>
<thead>
<tr>
<th>Providing Testing Locally or a Through A Centralised Service</th>
<th>Testing is centralised</th>
<th>Centralised testing may mean longer turnaround times for tests as they struggle to meet the demand from multiple hospitals</th>
</tr>
</thead>
<tbody>
<tr>
<td>Testing is localised</td>
<td></td>
<td>Localised testing may mean that testing is of poorer quality as pathologists do not get to see as many samples so have less opportunity for learning</td>
</tr>
<tr>
<td>Barriers to Introducing Future Barriers to Precision Medicine</td>
<td>Less accurate IHC screening</td>
<td>The IHC screening stain for ROS1 mutations has a lower specificity than that of ALK and so more tests will have to be sent for more costly and complex FISH testing</td>
</tr>
</tbody>
</table>

### 4.4.1 Managed Entry of Precision Medicine for NSCLC

Strategies to manage the entry of precision medicine for NSCLC, such as the process of technology appraisal by NICE, was an identified barrier to patients being able to access these test-treatment interventions. Some interviewees, mainly those from an oncology background, suggested that patients could previously access new medicines by enrolling in a trial or through early access to medicines (EAM) programmes. For *EGFR* and *ALK* targeted medicines, there was now clearly a gap between the availability of the medicine through a trial and EAM programmes and in standard clinical practice. This gap was the result of the appraisal process by NICE:
“but then there’s a delay between…with NICE approval or CDF (cancer drugs fund) approval and so it just means that there’s often windows of a few weeks to a small number of months where it’s not clear where you can get the stuff from” [C3, Oncologist, Town based hospital]

This interviewee points out that the length of this gap has been observed to be shortening in recent years:

“[NICE are] much more responsive than they were two or three years ago” [C1, Oncologist, City based teaching hospital]

This clinician believed that the observed responsiveness had been achieved by an improved process of “horizon scanning” and working with pharmaceutical companies to identify new examples of precision medicine that are currently being investigated in phase three (clinical development) clinical trials. However, the same clinician also warned that the increasing number of such interventions may have consequences for NICE and hence patient access to precision medicine for NSCLC:

“because one of the problems we know is that there are lots of things happening, there are very exciting new treatments that are demonstrating efficacy and so we just need to be able to make sure that NICE can keep up with all of the new approvals that are coming through which will be an issue” [C1, Oncologist, City based teaching hospital]

4.4.2 Funding Mechanisms and Service Commissioning

The most commonly discussed barrier across service providers was a lack of clarity in the commissioning process for precision medicine, with a particular focus on the availability of funding for testing. Due to the low availability of biomarker testing before and at the time of NICE approval for a precision medicine, it is common for the manufacturer of the medicine to provide initial (‘pump-priming’) funding to the NHS to help conduct testing. The process, while loss-leading for the manufacturer, allows clinicians to start using targeted medicines when testing does not exist in the health system. However, many participants highlighted the problems which occurred when companies withdrew funding for \textit{EGFR} and \textit{ALK} mutation testing before NHS funding for tests had been established:
“the pharma companies, of course, kind of, got wind to that [there was no existing testing] and started to fund the testing, either centrally or locally for a period of time, which then became time-limited. And at the end of that time-limited period, there was quite a lot of uncertainty about who would be paying for the tests, and where the bills should be sent” [C4, Oncologist, City based hospital, regarding EGFR testing]

Geographical inequity in access to testing arising from difficulties in receiving reimbursement for testing was a common theme in the interviews. While some service providers struggled to provide testing due to a lack of available funding, others used different strategies to ensure a continuity of test provision. Some clinicians believed that hospitals with bigger reputations could provide more funds for testing. In some cases research funding was used to fund testing until NHS funding was made available:

“We didn’t set things up that way in [City based Hospital], so we didn’t have that situation because actually we were paying for it from research funds anyway. So [City based Hospital] have not had that sudden removal of testing because we were paying for it from our own funds in the first place and just continued to do so” [C1, Oncologist, City based teaching hospital].

After the removal of pump-priming funds for testing by pharmaceutical companies, there was a perceived need for the NHS to provide a clear commissioning route through which hospitals would be reimbursed for the tests they conduct. However, one problem identified by this participant was that:

“NICE, kind of, recommends the drug and then don’t really make any provisions for the test” [C4, Oncologist, City based hospital]

While service providers are mandated to provide a medicine for patients with specific biomarkers, there is no requirement for them to actually provide the stratifying tests. In addition, no formal reimbursement mechanism exists should they decide to fund the test (70). This gap in the process for agreeing funding appears to have caused significant issues for each test for some service providers:

“And then there was the testing requirements for looking for EGFR mutations and ALK alterations, so that the main issue is that although these drugs became recommended for use, there
was very little clarity about who would pay, who would pick up funding for the testing” [C4, Oncologist, City based hospital]

In this scenario, test providers would effectively be supplying testing for free:

“It all comes down to the funding. You see what happens is, in our experience, like NICE would approve the drug, but at no point are laboratories or the managers being supported of how they can recoup the costs of the test. So, like say for ALK, we found it costs £50 a day by immunohistochemistry but it’s about £120 if you do it by FISH So, we got that result but it’s about £50 for that and we’ve been doing ALK for about three years now and we still have not been able to get the money reimbursed” [TP2, Biomedical Scientist, City based teaching hospital]

The lack of service commissioning arrangements seems to have resulted in the cost of the tests not being reimbursed by NHS England. There were, however, clear negative consequences for hospitals in NHS England. Without the cost of testing being reimbursed, it was perceived that the pathology laboratory performing testing would experience a financial loss. No interviewee stated that patients weren’t receiving tests because of a lack of reimbursement for testing:

“I have to do what’s right by the patient and follow the guidelines as they’re set and it’s for the labs to fight with the CCG’s [Clinical Commissioning Group] to try and clarify how they’re going to fund it, but I think in reality, they don’t have enough money and therefore they’re just going to get more in debt because there isn’t a way round it” [C3, Oncologist, Town based hospital]

For many interviewees, it was not clear who should take responsibility for the cost of testing given who accrues the benefits. This confusion was because precision medicine involves a complex intervention comprising a test-treat strategy and involves processes using many different specialities including pathology, genetics, and oncology. As such while specialities such as pathology and genetics may be covering the cost of testing, the benefits are realised almost entirely in oncology:

“we are paying huge amounts still for sending tests for PD-L1 testing which is actually not our test. In a way it is a test for oncologists” [TP3, Histopathologist, City based teaching hospital]
If testing helps to prevent patients from receiving an intervention that will not benefit them, as was
the case when $EGFR$ testing for gefitinib was introduced, the expense paid for by pathology or
genetics may even save money for oncology services. The complexity of the testing pathway and
the lack of clarity about test commissioning appeared to create conflict between the stakeholders in
the provision of the precision medicine:

“historically these sorts of things have perhaps been picked up in pathology departments but, you
know, the more and more tests we do, the less and less they are willing to, kind of, handle those
budgets and the more pushback there’s been” [C4, Oncologist, City based hospital]

“The actual doing of the work, it’s not been a problem at all and we’ve got a really good
relationship with the clinic and they’re very grateful that we reflex test but now the managers are
going. “who’s paying for this?”, “who’s approved this?”, that sort of thing” [TP2, Biomedical
Scientist, City based teaching hospital]

Some test providers who were interviewed were aware of relevant guidelines for the introduction
of precision medicine for NSCLC but for many there was still confusion about test reimbursement:

“So what I found was when the ALK testing was rolled out there is a tariff paid by NHS England
which covers the cost of the test” [TP1, Histopathologist, Suburban hospital]

“...you just jump through hoops and everybody wants to know who is going to be pay for this test.
So, we’re at the short-fall at the minute with ALK” [TP2, Biomedical Scientist, City based teaching
hospital]

Even where funding for the day-to-day costs of testing is available, some participants in this study,
particularly pathologists, struggle to find funding to validate the tests in their own laboratories.
This problem emerged with the introduction of immunohistochemistry screening for $ALK$:

“So the reagent for the ALK test was £50 so if you wanted to check, say, 100 cases that would be
£5,000 that needed to come from somewhere to validate it on before you could start running it
clinically. So there was no funding laid out by the NHS for that, you’d have to try and just find it
locally or source it within your lab itself” [TP1, Histopathologist, Suburban hospital]
For some test providers the barrier of a lack of funding for test validation was mitigated through the provision of grants by the manufacturers of the relevant precision medicine. This potentially mutually beneficial agreement helps laboratories to provide testing to patients while also ensuring the company increases the number of patients who are deemed eligible to receive the medicine. This solution was used by different providers for both ALK and PD-L1 testing. This interviewee gave the example for ALK mutation testing

“[Pharmaceutical Company] gave us a grant and because we were paying £120 to do the FISH and £50 for the immunohistochemistry, [pharmaceutical company] gave the Trust a grant to validate so we didn’t incur that cost to validate” [TP2, Biomedical Scientist, City based teaching hospital]

Other interviewees seem to believe that the NHS should be more involved in providing funding for test validation and development:

“...we’ve got one or two people who do work solely on development of tests. But, most departments don’t, I don’t think, have that. So, having more funds, or at least allocated time to do development of new tests, that’s the main, sort of, barrier that we have, I think” [TP4, Clinical Scientist, City based teaching hospital]

Failing to provide funding for genetic and pathology laboratories to validate their tests could lead to heterogeneity in the level of test validation and therefore quality of testing across England. This heterogeneity in test validation and quality in turn may result in patients missing out on treatments that would benefit them or receiving treatments which do not benefit them due to poor stratification. For example, one pathologist believed that other laboratories might not be as rigorous in assuring the quality of their test:

“I think it's all very variable in that if someone else in another pathology departments sets up the same assay they might not necessarily ask the questions that we're asking in exactly the same way” [TP1, Histopathologist, Suburban hospital]

4.4.3 Logistics of Organising Testing to Guide Treatment
Precision medicine for NSCLC involves the provision of diagnostic tests that then inform whether a targeted medicine is appropriate and the logistics of organising this process involves the collaboration of multiple health service disciplines. These complex interactions between different departments can be a barrier to implementation. For example, pathologists are involved near the beginning of the precision medicine pathways for: EGFR testing in preparing tumour samples, in ALK testing in conducting IHC screening to determine the need for follow up FISH testing, and in fully conducting PD-L1 testing. Clinicians and geneticists suggested that delays occurring in pathology laboratories can have knock on effects for the rest of the clinical pathway:

“So, as soon as we receive the sample, we can get the results back very fast. But, sometimes, there’s a delay in receiving the actual sample from pathology, so that can delay things quite a bit sometimes” [TP4, Clinical Scientist, City based teaching hospital, regarding EGFR mutation testing]

“there’s a number of delays in that pathway. Partly it’s the pathologists remembering to do it, then the sample actually being physically sent, then the sample being received, then processed, then the result coming back” [C4, Oncologist, City based hospital, regarding ALK testing]

An additional barrier, which was highlighted by some participants, can be that the quality of the tumour sample is not sufficient for geneticists to conduct mutation testing:

“Quite often. It’s not very rare, actually. Or, it might even be that we do receive the sample on time, but the quality of the sample is not good enough for testing” [TP4, Clinical Scientist, City based teaching hospital]

These delays may be as a result of processing by pathologists or simply because it is difficult to extract a sufficient volume of tumour from patients who have advanced lung cancer (208). In the former case, additional samples can be requested from those kept in storage or through re-biopsy. In the case of the latter, treatment options may be limited:

“it can happen that either the patient refuses to have the second biopsy or the patient can’t wait to have the second biopsy, because of deteriorating clinical conditions and it’s better to start chemotherapy instead of waiting for another biopsy” [C2, Oncologist, City based hospital serving rural community]
Potential barriers in the quality and speed of service provided by pathology laboratories may occur due to an increasing pressure on services due to the volume of samples being processed and the increased workload involved in processing each sample. For example, pathologists first conduct haematoxylin and eosin stains which help to highlight abnormal cell structures from which a cancer diagnosis can be made. They will then determine the type of lung cancer present by looking at the cells and potentially conducting additional immunohistochemistry tests. If the cancer is found to be NSCLC, they will then prepare part of the sample to send for EGFR mutation testing and potentially conduct a number of IHC tests on the remaining parts of the sample for other markers such as ALK or PD-L1.

An additional barrier presented in providing PD-L1 testing is the quantitative rather than qualitative nature of the test. This means that there is some subjectivity involved in the interpretation and reporting of the results of IHC stains:

“if there's a little bit of tumour in a big resection slice like that might be a resection slice and you might have thousands of tumour cells in that. And then they're asking you to say, well, is it one per cent of the tumour positive or less than one per cent? So if you have a little bit of staining in the tumour but it's less than one per cent it will be taken as a negative result, but if it's one per cent it's taken as positive for second line therapy. So that can be really hard to decide sometimes” [TP1, Histopathologist, Suburban hospital]

The tumour samples sent for different tests often need to be prepared in different ways, adding an additional layer of complexity and a potential barrier to providing prompt results:

“What we do in our lab is cut sections and send the slides, but for tests like PD-L1 they want the whole block because they want to cut it fresh so that interpretation is not missed, not changed. So, we all needed to understand how we can do it best” [TP3, Histopathologist, City based teaching hospital]

Despite the increased workload, pathologists suggested that they struggled to find sufficient resources to support their services, therefore creating another barrier to returning test results promptly and allowing treatment to be started quickly. This problem does not seem to have been faced by genetics labs.
“that’s the major problem we are facing in our department because as we know, budget is available for PD-L1 testing and I get every time the pharmaceutical companies com[ing] to me. I say I can understand that, but we don’t have enough people so we just need to see how much we can do” [TP3, Histopathologist, City based teaching hospital]

“the genetics lab had the capacity there, they’ve had to increase their capacity to incorporate other tests but they had capacity for that test at the time they implemented it” [C5, Oncologist, City based teaching hospital, regarding EGFR mutation testing]

The increasing number of different tests that are now involved in the pathway for patients with advanced NSCLC can act as a barrier to patients receiving prompt access to a relevant targeted medicine as delays to processing one test can cause delays in conducting and receiving results for others. As the mutation tests are currently done in sequence, it is common to wait for the results of EGFR or ALK before sending more of a tumour sample for PD-L1 testing:

“sometimes in the past what has happened is we have received a request for EGFR and ALK and then we’ll get a request, oh, can you do PD-L1? And by that time the sample has already gone for testing for ALK in a different place. So, we’ll have to ask for a block and then it causes time delays” [TP3, Histopathologist, City based teaching hospital]

In addition to the complexity caused by sequential testing, the different tests are often sent to different laboratories and potentially even different trusts for analysis:

“the additional issue is that our pathologists have to collect the sample and send part of it to Sheffield to perform the EGFR and ALK, and part of it to Birmingham. So two different places. And I understand an additional complication for the pathologists, because I think they want to wait and see if there is enough tissue for EGFR and ALK. And once they know there is enough tissue, then they send the request for the PD-L1” [C2, Oncologist, City based hospital serving rural community]

4.4.4 Providing Testing Locally or a Through A Centralised Service
The size and location of the laboratory conducting ALK IHC or FISH testing was identified as a potential barrier to the timely use of precision medicine. However, there were significant differences in participant attitudes to whether testing should be conducted more locally or should be conducted in fewer, centralised laboratories. In addition, opinions on the issue of centralisation or localisation appeared to depend on the nature of the test and role of the service provider. For example, it was generally considered that as EGFR testing required specialist knowledge this test would need to be conducted in larger genomic medicine centres:

“Yeah, yeah. I think EGFR can’t be established in all places. It has to be where molecular tests can be done” [TP3, Histopathologist, City based teaching hospital]

There was substantially more disagreement about where ALK testing, rather than EGFR testing, should be conducted. While ALK testing was originally conducted using centralised FISH (fluorescence in situ hybridisation)-based genetic testing, the move towards immunohistochemical screening for the mutation increased calls for the analysis to be moved into local pathology labs. Many participants believed that moving testing “in-house”, often in the same hospital that the patient was being treated at, would mean cheaper testing and a shorter turnaround time for tests leading to better outcomes for patients. This was seen to be particularly true of PD-L1 testing, another immunohistochemical-based test, which is currently only conducted by a small number of centralised pathology labs.

“So it’s a problem at the moment and when that comes in house in the next few weeks it will take away that terrible time delay that will take away the uncertainty of me having to say to the patient “well actually I’m not sure of what you’re initial treatments going to be”” [C1, Oncologist, City based teaching hospital]

“So, we’re getting good correlation, so I think we should be able to finish our validation in the next two weeks or so and fingers crossed, if we can get the money, we can do it in half because the turnaround time in-house is going to be a day, whereas with Liverpool currently it’s taking about two weeks ” [TP2, Biomedical Scientist, City based teaching hospital]

Some test providers were more sceptical of the drive to localise pathology testing. Some pathologists believed that policy makers were pushing for local testing because the large
laboratories did not have the capacity to meet demand. Furthermore, these individuals highlighted the fact that they were being asked to set up testing without any funding being available.

“I find there's this huge pressure and bottleneck from the whole structure of the trust within the NHS because there's a drive to save money and they don't actually want you to set up a new test locally if they think another hospital can do it because they want to just save money within the trust” [TP1, Histopathologist, Suburban hospital]

“I think the main problem which the bigger hospitals are now finding is the turnaround time. They can’t deal with it. They have the pressure of cases, so it’s the balance between expertise and turnaround time” [TP3, Histopathologist, City based teaching hospital]

In addition to a trade-off between cost and turnaround time for IHC testing, questions were raised about whether localising testing would reduce quality. This is because large laboratories processing large volumes of tests would be able to learn from experience while small laboratories might be unfamiliar with test methodology.

“we’ll have more local testing although it won’t necessarily be done in every single lab because the pathologist needs to have a number of tests that they do on a regular basis in order to be confident in the scoring” [C1, Oncologist, City based teaching hospital]

4.4.5 Barriers to Introducing Future Examples of Precision medicine in NSCLC

Participants identified a number of new developments in the provision of precision medicine for patients with NSCLC as well as the potential barriers that were faced to support their implementation. During the time-scale of the interview process for this study, NICE approved the use of the drug crizotinib for patients with ROS1 mutations. When asked about potential barriers to introducing testing for ROS1 mutations, there was some optimism that implementation would be smoother than for previous tests. This perceived easier implementation was because many laboratories have set up ROS1 testing for research purposes and as the testing process is similar to that of ALK testing, because it also uses an immunohistochemical-based test prior to FISH testing. This interviewee spoke about establishing ROS1 testing before the approval of crizotinib:
“[Tertiary hospital] have the assay but we just do that, we just do ROS1 on some of the adenos (adenocarcinomas), basically give ROS1 testing at [tertiary hospital] as an immunochemical test. And then if they're positive they'd get FISHed the same way you would for ALK.” [C1, Oncologist, City based teaching hospital]

There were still deemed to be some potential barriers to introducing ROS1 testing. Again, a lack of clear commissioning routes for the test was highlighted and one participant highlighted that the IHC screen for ROS1 was not as good as that for ALK meaning that more samples sent for the more expensive FISH test would be negative for the mutation.

“But I think ALK is the most straightforward interpretation of all of them. ROS1 will have more cases which will go for FISH so I think we’ll have to be careful about the funding and budget about those should be there so, yeah” [TP3, Histopathologist, City based teaching hospital]

Given the increasing number of tests available to stratify treatment for patients with NSCLC and the current approach of running tests in sequence, there is a risk of delaying the start of treatment for patients while waiting for test results. However, a number of developments may help to reduce this risk by ensuring that some or all of the tests are conducted at the same time. Some genomics laboratories have begun offering panel tests in which all potentially relevant mutation tests are conducted at the same time. However, such panel tests can be technically difficult to set up as the methods for testing for EGFR and ALK mutations are different. A further development, next generation sequencing (NGS), may counter this issue by allowing all mutations to be searched for using a single test rather than a group of individual tests:

“And ideally you need a panel that includes...that can cover the translocations as well and then you could actually test for ALK and ROS1 in the panel and you could have one test, more or less, except maybe PD-L1 as well, that's separate. Yeah, but I think that's what needs to be aimed for, that you just make one test, that just makes it easier” [TP1, Histopathologist, Suburban hospital]

Currently in the UK, NGS-based testing is used only in research but not in clinical practice due to its high cost. However, as more mutations become relevant in clinical decision making, the balance between the cost of NGS, and associated interpretation of the result, and the cost and interpretation of a number of individual tests may become more favourable. NGS-based testing also has the advantage of requiring less tissue from patients and may reduce the number of repeat biopsies that
are required (209). However, as highlighted by pathologist 1 in the above quote, NGS-based testing would not be able to guide treatment of PD-L1 targeting therapies as currently only immunohistochemical-based tests are available for PD-L1 overexpression.

In addition, as patients’ tumour profiles can change over time additional tests may be required regardless of how advanced the additional test methods were. Such changes include the addition of resistance mutations such as EGFR T790M (210) and ALK G1269A (211) following treatment with EGFR TKIs and ALK inhibitors respectively. To identify these potential mutations in patients who stop responding to treatment, additional testing is required which in the past has required an additional invasive biopsy to be conducted. Recently, circulating tumour DNA (ctDNA) testing has been introduced by some laboratories which allows such mutations to be identified from a blood sample rather than a tumour sample (212). Some participants suggested benefits to using blood based tests instead of tumour samples for initial identification of mutations:

“If anything, I think, I mean, obviously if you’d skipped the whole pathology step, then that would save a lot of money. Especially if you don’t do the biopsies, but the actual test, the cost is very similar.” [TP4, Clinical Scientist, City based teaching hospital]

“sometimes they’re [the patient] too sick, sometimes you have multiple failed biopsies, sometimes you...yes so, we had a patient that we tried to repeatedly biopsy and we just couldn’t get enough tissue to do it and we did a serum test and it was positive and that was our answer” [C3, Oncologist, Town based hospital].

4.5 Discussion

This qualitative interview study identified 17 barriers that may impede the introduction of examples of precision medicine for NSCLC. These barriers were grouped into five themes: the managed entry of precision medicine for NSCLC; the commissioning and reimbursement of precision medicine for NSCLC and specifically the test component of precision medicine; the complexity of the logistics around providing tests; opinions about whether test provision should be localised or centralised; and opinions about future developments, including potential barriers to their introduction, in precision medicine for NSCLC.
The existence of these barriers may explain the slow adoption of the test-treat interventions into clinical practice for treating patients with NSCLC. Therefore, these barriers can also be identified as capacity constraints. The development of this typology of barriers, and associated capacity constraints, to the introduction of examples of precision medicine for NSCLC may aid in predicting the potential barriers that may be faced in introducing future examples of precision medicine in NSCLC, specifically, and other disease areas, generally. By being aware of the barriers and taking action to address them before the introduction of a new test-treat intervention, it may be possible to increase the speed of implementation of precision medicine into clinical practice and ensure that all patients receive appropriate, high quality testing and treatment.

This qualitative study identified capacity constraints relevant to test-treat strategies for NSCLC. Some of these constraints are specific to this disease area and did not appear in the systematic review reported in chapter 2 or in the broader literature. For example, a significant barrier identified in this study was the lack of a clear pathway for reimbursing pathology laboratories for conducting testing. An additional problem which may be specific to the area of precision medicine for NSCLC is delays associated with sequential testing given the volume of biomarker targeted treatments that are approved for use for this condition. However, there were some similarities with the capacity constraints to introducing examples of precision medicine more generally which were identified in the systematic review reported in chapter 2. For example, in the systematic review (see section 2.3.4), the cost of testing or treatment was seen as a significant constraint to introducing examples of precision medicine. In the qualitative interview study the cost of validating testing and issues around test reimbursement were frequently discussed by participants as relevant capacity constraints. The quality of testing barrier identified in the systematic review (see section 2.3.5) bears some similarities with the points raised in the interviews around the effects of localising or centralising testing and the delayed turnaround time from sequential testing. The participants also mentioned delays caused by NICE in evaluating new examples of precision medicine that aligned with the barrier regarding the need for clinical and economic evidence to address decision uncertainty identified in the systematic review (see section 2.3.7).

Some of the specific barriers raised in this qualitative study are consistent with barriers mentioned in reports and papers authored by institutions involved in the provision of examples of precision medicine. For example, a recent report by Cancer Research UK highlighted the strain on pathology laboratories in the NHS (213). In particular, this report highlighted increasing workloads for cytopathologists who diagnose cancer while the number of new professionals working in those
roles was growing at a slower rate. The report suggested that while molecular pathologists who are involved in biomarker tests were currently able to manage their workloads, this was partly due to the current underuse of these tests to guide treatment with precision medicine.

The potential barriers associated with localised testing have been recognised more broadly across pathology testing and there is currently a strategy to create a national network of pathology laboratory hubs (214). It has been argued that “consolidating pathology services allows for the most consistent, clinically appropriate turnaround times, ensuring the right test is available at the right time” (215). Furthermore, it has been suggested that centralising testing could save the NHS £200million by 2020-2021 through economies of scale. The potential to benefit from economies of scale when centralising testing has also been highlighted by Buckell et al. (2015) who suggest potential efficiency savings of 13% or up to £390 million per year (216).

The extent to which the findings of this study generalise beyond the disease area of NSCLC and the setting of the NHS is unclear. Many examples of precision medicine share the potentially complex logistics involved in using one or more tests to guide treatment and given that many of the barriers occur in the testing pathway these may also apply for other examples of precision medicine. For example, delays caused by the NICE evaluation process may potentially occur for any treatment where the cost-effectiveness of the intervention needs to be evaluated. Alternatively, problems with using one block of tissue to conduct many different tests are likely to be specific to the area of NSCLC which has many approved targeted treatments. Many of the barriers associated with the technical and logistical aspects of pathology testing may also be experienced in different health systems while barriers related to financial reimbursement and the capacity of pathology laboratories in terms of staffing levels are likely to be more specific to the health system. Given that many of the barriers will closely depend on the context of the health system, studies investigating how to effectively implement new examples of precision medicine should arguably focus on these specific health system contexts and involve stakeholders from these systems in the research.

There were some limitations to this qualitative study. It transpired during interviewing that one of the targeted samples in this study, the service commissioners responsible for financing for tests and targeted treatments, were not identifiable as individuals. This specific role is not part of the NHS service commissioning processes. One service commissioner was interviewed, but their primary commissioning role was in the provision of germline genetic and genomic tests. The individual
only had second-hand knowledge of the commissioning of the companion diagnostic tests required for precision medicine in NSCLC. It became apparent through this participant, and those working in other specialities, that in essence the role of service commissioner for such companion diagnostics does not exist and that commissioning arrangements were often laboriously discovered following conflict between hospital managers and test providers.

There were also difficulties in interviewing participants in the test provider sample. Due to a national policy of restructuring NHS genetics and genomics services during the recruitment phase for this interview, it was difficult to recruit clinical geneticists who make up a sub-section of the test provider sample due to their excessive workload. It was possible that due to the small number of these participants who were interviewed that saturation was not reached in this sub-sample. This may be likely as new themes had been identified in the transcripts of the most recently recruited pathologists and geneticists. Further recruitment in these samples could have resulted in additional barriers to the introduction of examples of precision medicine being identified. The omission of key barriers may have been likely given that the majority of the barriers identified in this study originated in the testing pathway.

It was also difficult to identify participants who had experienced the introduction of *EGFR* mutation testing and treatment with gefitinib. At the time of recruitment it had been over 8 years since this example of precision medicine had been approved by NICE and few participants were working in an NHS role at this time. As such, the coverage of the barriers to the introduction of *EGFR* mutation and targeted treatment appeared, in this study, to be less extensive than for *ALK* or PD-L1 testing. This extensive length of time also created a greater risk of recall bias when compared to *ALK* testing (approved four years prior to the study), and PD-L1 testing (approved two years prior to the study but still being implemented). However, the extent and range of barriers that were identified in this study was sufficient for generating a typology of capacity constraints and for informing a subsequent quantitative analysis of the impact of capacity constraints on the cost-effectiveness of introducing a test-treat strategy for treating patients with NSCLC (see Chapter 6).

This study has focussed on the experiences of stakeholders involved with introducing existing examples of precision medicine. However, in applying the methods described in this thesis to account for the impact of capacity constraints, researchers will likely have to predict the potential barriers that will be faced to introducing a specific example of precision medicine in the future. To
elicit such predictions of the future from stakeholders, more complex qualitative methods may be required and these may be drawn from the field of implementation science. Implementation science is “the scientific study of methods to promote the systematic uptake of research findings and other evidence-based practices into routine practice, and, hence, to improve the quality and effectiveness of health services and care” (217).

A number of theories and frameworks have been developed in the discipline to aid in the identification of key factors which may aid or impede the implementation of an intervention (218). For example, normalisation process theory addresses the key research question of “How can those factors that promote or inhibit the normalization of complex interventions be identified, conceptualized, and evaluated?” (219). Normalisation process theory outlines a series of actions which are required in order for an intervention or policy to be normalised in clinical practice, with these actions falling within four key constructs: coherence; cognitive participation; collective actions; and reflexive monitoring (219–221). Coherence relates to the actions needed for organisation and individuals to make sense of a new intervention, cognitive participation relates with how people engage with the act of implementing an intervention, collective action is how people enact implementation, and reflexive monitoring is the appraisal of implementation (220). Each of these core constructs contains four more specific sub-domains which further describe the actions needed for implementation.

An alternative, or complimentary, approach to understanding the conditions needed for the effective implementation of an example of precision medicine is the development of a theory of change (222). This approach involves exploring with stakeholders the mechanisms by which a new intervention or policy creates improvements in outcomes for patients and the conditions necessary for these improvements to be realised. In fact “determining what contextual conditions are necessary to achieve the outcomes, what resources are required to implement the interventions, and how the program gains the commitment of those resources are crucial outputs of the process” (222). Using one of these theories, or other implementation frameworks, such as the Consolidated Framework for Implementation Research (218), as a template for a semi-structured interview schedule or other qualitative methods may help to guide participants towards thinking about a wide range of barriers to implementing a new intervention.

4.6 Conclusion
This qualitative study, that used semi-structured telephone interviews, has shown that significant barriers exist to the introduction of examples of precision medicine in NSCLC. These barriers may result in patients not receiving access to potential beneficial testing and treatment, healthcare providers offering testing at a financial loss, and poorer quality testing leading to worse outcomes for patients. While some progress in addressing these barriers has been made, some have been faced in the introduction of all example of precision medicine currently available in the NHS.

4.7 Summary of Chapter 4

Chapter 4 identified 17 barriers that have been faced in introducing examples of precision medicine for non-small cell lung cancer. These barriers were rarely mentioned in published technology appraisals of the interventions.

A key element is being able to describe the capacity constraints relevant to the implementation of precision medicine. It is also necessary to be able to quantify the impact of capacity constraints. The barriers to precision medicine in NSCLC identified in this qualitative study will be used as capacity constraints in two empirical studies that use quantitative methods. Chapter 5 will present a case study economic evaluation of an example of precision medicine for NSCLC which was conducted in line with current practice. This will provide baseline estimates of the cost-effectiveness of an example of precision medicine in the absence of capacity constraints.

Chapter 6 will outline the identification of key capacity constraints to these interventions based on the findings of this study. It will then seek to quantify the impact of the selected capacity constraints on the cost-effectiveness and total net benefit of the example of precision medicine using the static value of implementation method developed in chapter 3.

Chapter 7 will explore how the dynamic value of implementation method developed in chapter 3 can be used to value three hypothetical investments in improving health system capacity to facilitate the introduction of the example of precision medicine.
Chapter 5

Generating the Decision-Analytic Model-Based Cost-Effectiveness Analysis of an Example of Precision Medicine: a Case Study

5.1 Chapter 5 Summary

The results of the empirical study described in chapter 4 identified relevant capacity constraints that may influence the estimated cost-effectiveness of precision medicine.

Chapter 5 reports the replication of a published decision-analytic model to estimate the incremental costs, quality adjusted life years and cost-effectiveness of an example of precision medicine for NSCLC. This published decision-analytic model was submitted by a manufacturer to NICE as part of a technology appraisal. This decision-analytic model will provide the baseline result to serve as a comparator for subsequent results produced by a model adapted to incorporate capacity constraints using static methods (see Chapter 6) and dynamic methods for estimating the value of implementation (see Chapter 7).

5.2 Introduction

A rising number of examples of precision medicine are available for clinicians to prescribe treatments for patients with NSCLC. These include treatments that target \textit{EGFR} mutations, \textit{ALK} and \textit{ROS1} alterations, and PD-L1 overexpression among other biomarkers (see also section 4.2). These treatments that target specific tumour mutations often come at a high cost. For example, erlotinib for patients with \textit{EGFR} mutations currently costs £1,631.53 per patient month. Nivolumab to target PD-L1 overexpression costs up to £2,633 per patient per month. The cost of these medicines often means that while they may offer health benefits to patients, they are more expensive than current treatment options. This means that funding to provide the targeted medicines must be taken from other areas of the health system. Part of the role of the National Institute for Health and Care excellence (NICE) is to evaluate such high cost medicines and to determine whether the benefits they offer outweigh their additional cost in a technology appraisal (44).
As part of NICE single technology appraisal of a new medicine, produced by a single manufacturer, the process requires the pharmaceutical company to submit a decision-analytic model-based cost-effectiveness analysis of the intervention under appraisal (44). This decision-analytic model-based cost-effectiveness analysis is then critiqued by an Evidence Review Group (ERG) commissioned by NICE with the result that the manufacturer is often requested to make changes to their assumptions to produce new estimates of the cost-effectiveness of their intervention.

The economic evaluations submitted to NICE and appraised by a committee of experts, the Technology Appraisal committee, are conducted to standards set out in the NICE reference case (44). The NICE reference case requires that manufacturers report on the potential for resource constraints to impact on the implementation of a new intervention. However, there is no requirement to incorporate the impact of such constraints in the decision-analytic model-based cost-effectiveness analysis. Some published economic evaluations have incorporated capacity constraints and shown that these could impact on the cost-effectiveness of precision medicine (see Chapter 2, section 2.3.9 and 2.3.10). There are many potential barriers to the introduction of precision medicine for NSCLC (see Chapter 4, section 4.3.1 to 4.3.5). In many cases, these barriers mean that service providers struggled to implement precision medicine after the test-treat strategy had been approved by NICE.

5.2.1 Aims and Objectives

The aim of this study was to replicate a published decision-analytic model-based cost-effectiveness analysis of an example of precision medicine in lung cancer. The replica decision-analytic model was created to provide baseline estimates of the cost-effectiveness of a test-treat strategy for NSCLC in the absence of health system capacity constraints.

In order to achieve this aim, this study had four objectives to:

- select a case study to use as an example of precision medicine for treating patients with NSCLC;
- replicate a published decision-analytic model-based cost-effectiveness analysis of the selected case study;
- produce a set of baseline results representing the cost-effectiveness of the selected case study in the absence of capacity constraints.
• To compare these results with the results produced by the baseline model to those produced by the model it sought to replicate

5.3 Methods

This study used decision-analytic modelling methods conducted in line with published guidance (49,223) and reported in line with published guidance (120). A published decision-analytic model-based cost-effectiveness analysis of a selected case study was replicated, but with some necessary adaptations, using the software programme R (224).

5.3.1 Selecting the Case Study

A case study is an “intensive (qualitative or quantitative) analysis of a single unit or a small number of units (the cases), where the researcher’s goal is to understand a larger class of similar units (a population of cases)” (225). In this thesis, one key aim is to determine the potential impact of the inclusion of capacity constraints on the relative cost-effectiveness and net societal benefit of examples of precision medicine. A case study was required to illustrate the potential impact of capacity constraints while allowing the potential generalisation of the findings to the broader area of precision medicine.

The selected case study was required to possess two key characteristics to provide the potential for generalisation of findings to precision medicine more generally. The selected case study must be representative of the broader population of examples of precision medicine (225). A systematic review (see Chapter 2) identified that there were a diverse range of interventions which could be deemed to represent “precision medicine”. To select a relevant case study, a consistent focus was taken in this thesis, and precision medicine was interpreted to be as a single test-treat intervention (see Chapter 1). Using this focus may limit the generalisability of findings to interventions where a test is provided but the results do not lead to any change in treatment. For example, genetic testing for the inherited eye conditions collectively known as retinitis pigmentosa may help to inform the prognosis and management of the condition but there is currently no treatment available (226).

The further consideration for selecting a relevant case study was that it should exhibit variation in “the dimensions of theoretical interest” (225). This consideration implies that the included case
study should provide the opportunity to understand the impact of a range of capacity constraints with different potential mechanisms of action on influencing the relative cost-effectiveness of an example of precision medicine.

A cross cutting theme in the identified barriers to introducing examples of precision medicine, both generally (101,161) and for patients with non-small cell lung cancer (NSCLC) specifically (see Chapter 4), was that such barriers most commonly arose from the testing component of the test-treat intervention. Therefore, the selected case study of precision medicine was primarily chosen on the characteristics of the test component of the test-treat intervention. The case study was chosen with the aim of being representative of emerging examples of precision medicine, in general, and for NSCLC, specifically while also ensuring ‘sufficient’ diversity in the testing process to demonstrate the potential impact of a range of different defined capacity constraints.

The selected case study used in this thesis was anaplastic lymphoma kinase (ALK) testing that uses immunohistochemistry (IHC) and fluorescence in situ hybridisation (FISH) to target treatment with an appropriate treatment such as an ALK inhibitor (crizotinib). This case study was a NICE Single Technology Appraisal and published as part of NICE T296 (227). The relevant comparator for this test-treat intervention was no testing and a chemotherapy agent (docetaxel). ALK testing involves an IHC stain test, conducted by a pathologist, to first screen for patients with potential ALK mutations (228). The IHC-based test has a reported high specificity (95%) meaning that few (5%) patients with ALK mutations will be given a negative result that is not correct (low proportion of false-negative results) (229). The sensitivity of the IHC-test is not as accurate (90%) meaning that there is a ~10% chance that patients will be given a false-positive result. This relatively low sensitivity means that a confirmatory, more accurate, genetic-based fluorescence in situ hybridisation (FISH) test conducted by a molecular geneticists is needed to rule out potential false-positive results (229). The use of the IHC-test as a ‘screen’ will reduce the number of the more expensive genetic-based FISH-test required to achieve a confirmatory result of ALK mutation status (230).

5.3.2 Construction of the Decision-Analytic Model-Based Cost-Effectiveness Analysis

This section outlines the replication of a published decision-analytic model to estimate the cost-effectiveness of ALK testing using an IHC-test and FISH-test to identify the eligible patient
population for treatment with an \textit{ALK} inhibitor (crizotinib) compared with no testing and docetaxel (227).

\textbf{5.3.3 The Decision Problem}

The decision problem was to identify the incremental NHS costs and QALYS for \textit{ALK} mutation testing using an IHC-test and FISH-test to guide treatment with 250 mg crizotinib twice daily for patients with previously treated non-small cell lung cancer (NSCLC) compared with no testing and universal treatment with 75mg/m$^2$ docetaxel. This decision problem differs to that used to inform the development of the source model that sought to estimate the cost-effectiveness of 250mg crizotinib twice daily compared to 75mg/m$^2$ docetaxel for non-small cell lung cancer patients who had already received a positive \textit{ALK} mutation test. In this study, a decision tree was added to the source model to represent the testing component of the intervention (see section 5.3.9) as most of the barriers to introducing \textit{ALK} testing and targeted crizotinib or docetaxel lay in the testing pathway (see section 4.4).

\textbf{5.3.4 Study Perspective}

The costs and consequences of a health care intervention and comparator can be experienced by a wide range of individuals and organisations including patients, the health care service, the social care service and other public sectors. The study perspective of an economic evaluation is defined as the selection of the groups whose costs and consequences will be included in the selected analysis, such as cost-effectiveness analysis (231,232). The relevant study perspective for the stated decision problem assumed the health care service perspective and is consistent with the perspective recommended in the NICE reference case (44).

\textbf{5.3.5 Time Horizon}

The time horizon in an economic evaluation is defined as the duration of time over which changes in costs and consequences associated with the specified intervention relevant comparators are assessed (232,233). This case study used a time horizon of 15-years as only 0.01% of patients with NSCLC are expected to live longer than this (234).
5.3.6 Study Population

The study population is the group of patients for whom the intervention is being appraised (44). The cost-effectiveness of an intervention may differ depending on which groups of patients receive the intervention. The relevant study population for this case study was all patients with confirmed NSCLC who had previously tested negative for epidermal growth factor receptor (EGFR) mutations. This restriction of the relevant study population was because testing for treatment selection in NSCLC are conducted in sequence with the more common EGFR mutation being tested for first in the test-treat pathway (see Chapter 4). In addition, the presentation of ALK and EGFR mutations are generally mutually exclusive so patients with EGFR positive NSCLC are unlikely to also have ALK positive cancer (235,236). In Duarte et al., (2014) the population of patients in the UK with NSCLC estimated by the manufacturer of crizotinib was stated to be 19,561. Given that approximately 15% of patients with NSCLC have EGFR mutation positive tumours in predominantly Caucasian European countries such as the UK (237), the study population who would receive ALK testing is approximately 16,627.

5.3.7 Intervention and Comparator

The relevant intervention for this analysis was defined as testing for ALK mutation using a combination of an IHC-test as a screen followed by FISH-test to identify patients with NSCLC eligible for crizotinib. Patients not eligible for crizotinib were assumed to be eligible for docetaxel at a dose of 75mg/m². Patients with positive IHC tests receive FISH testing and those with positive results receive 250mg of crizotinib daily until disease progression while those with negative results receive standard chemotherapy. In this case study this general chemotherapy agent was a 75mg/m² dose of docetaxel.

The comparator for this analysis was no ALK mutation testing and universal chemotherapy with docetaxel in 75mg/m² doses. Docetaxel and pemetrexed were considered as comparators but while they exhibit similar outcomes for patients, at the time of appraisal (2014) pemetrexed was used as a first line therapy and docetaxel was used as a second line treatment (234,238). As crizotinib was evaluated as a second line treatment, docetaxel was therefore deemed the most appropriate comparator.
5.3.8 Model Conceptualisation

Decision-analytic models provide a framework for comparing the costs and outcomes of an intervention and comparator using a simplified version of reality in order to “inform medical decision and health-related resource allocation questions” (239). The process of conceptualisation involves converting the clinical experience of treatment with the intervention or comparator into this simplified version of reality. The approach to conceptualisation used in this study was to reproduce an existing example of a published decision-analytic model-based cost-effectiveness analysis of ALK testing and targeted treatment with crizotinib or docetaxel (234).

The source model for this study was the decision-analytic model submitted by the pharmaceutical company Pfizer. A description of the original decision-analytic model, produced by the manufacturer, was identified in the published documents that supported NICE TA296 (234). A fully executable version of this model is not in the public domain. So, it was necessary to conceptualise and build the decision-analytic model de novo. This was done by using the published critique of the Evidence Review Group (234). Evidence Review Groups are independent academic groups who are asked by NICE to critique the evidence of effectiveness and cost-effectiveness submitted by the manufacturers of interventions to NICE (240). The Evidence Review Group often suggest changes to the parameters used in the models to reflect different assumptions regarding the effectiveness, costs and outcomes of the model. This in turn leads to a number of potential cost-effectiveness estimates being presented. However, in this case study, the original submitted model was used as the base line analysis.

The manufacturer submission features an economic evaluation which mainly focuses on the medicine component of test-treat intervention. This submission provided the key parameters needed for the base line case study model. However, the benefit of precision medicine is derived from the information provided by the testing component in effectively stratifying patients to the treatment which will benefit them most. As only minimal details of the testing component were provided in the documentation supporting NICE TA296, additional information was taken from relevant published economic evaluations identified in the systematic review reported in chapter 2 (128,241,242).

5.3.9 Model Structure
There are a number of types of decision-analytic models which are available as a framework for conducting economic evaluations of healthcare interventions (48). Each type of decision-analytic model has strengths and weaknesses to address specific decision problems. The selection of the appropriate decision-analytic model type is critical to best representing the experiences of patients in receiving the intervention of interest. This case study used a cohort-based state transition Markov model in keeping with the original published model to be replicated (234). State transition Markov models are widely used in economic evaluations of treatments for cancer as they allow the duration that patients are expected to stay in a given health state to be modelled and for this to impact on health outcomes and costs. This allows the key outcomes of cancer trials, progression free survival (PFS) and overall survival (OS), to be used in the model.

The source model was a three health-state transition Markov model that represents the experiences of patients living with NSCLC. All of the patients receiving either the intervention of comparator start in a progression free health state. Patients can then transition to a progressive disease health state based on the expected progression free survival (PFS) of the medicine they are receiving. Patients can also move from the progression free or progressive disease states to death depending on the expected overall survival (OS) of the medicine they are receiving. The PFS and OS for crizotinib and the PFS for docetaxel were derived from reproduced Kaplan-Meir (KM) curves from clinical trials while the estimate of OS for docetaxel was derived by applying a hazard ratio to the KM curve for crizotinib (see table 5.2). The differences in PFS and OS between crizotinib and docetaxel therefore drive differences in the movement of the patient cohort through the model. The treatment model had a cycle length of one-month and ran for 15 years resulting in a total of 180 cycles.

The source model for the case study did not feature a representation of the testing process. As most capacity constraints originate in the testing pathway, it was important that the model was adapted to represent the test pathway in the model. This was achieved by structuring a decision tree to link with the replicated state transition Markov model. The construction of the decision tree was based upon published economic evaluations of ALK testing and targeted treatments identified in the systematic review reported in chapter 2 (128,241,242). In the decision tree, the eligible population of patients were assumed to be offered testing or no testing. Those not offered testing were assumed to all start docetaxel chemotherapy. Patients receiving testing first receive an IHC test for ALK mutations. If this test result is negative they then receive docetaxel. If this test result is positive then they go on to receive FISH testing. Patients then receiving a positive FISH test results
for *ALK* mutations receive crizotinib treatment while those receiving a negative result receive chemotherapy.

**Figure 5.1: Decision tree**

In this decision tree the rectangular box represents the decision problem, the circle icons represent chance nodes, and the diamonds represent entering into the state transition Markov model (see figure 5.2).

Abbreviations: IHC=immunohistochemistry; *ALK*=anaplastic lymphoma kinase; FISH= fluorescent in-situ hybridisation.

---

1In this decision tree the rectangular box represents the decision problem, the circle icons represent chance nodes, and the diamonds represent entering into the state transition Markov model (see figure 5.2).

Abbreviations: IHC=immunohistochemistry; *ALK*=anaplastic lymphoma kinase; FISH= fluorescent in-situ hybridisation.
The decision-analytic model was constructed using the programming language R (version 3.53) (224). R is an extremely flexible approach to constructing models which allows various mathematical analysis techniques to be combined. This flexibility made it useful for linking different decision analytic modelling types and for subsequently building in capacity constraints into these models (see Chapters 6 and 7).

The construction of the decision tree and state transition Markov model was made simpler by the use of pre-existing user packages in R for various aspects of the analysis. The data.tree package in R was used to formulate the testing pathways and to calculate the conditional probabilities and costs of patients receiving different test results (243). The heemod package in R was used to produce the state transition Markov model (244). The coding of the heemod package compares the costs and QALYs of a single cohort receiving either the comparator or intervention and then computes the total costs and benefits of both arms and calculates the ICERs. The package therefore assumes that the same number of patients receive the intervention and comparator. However, in both case studies the “treatment” arm involved some patients receiving the precision medicine and some patients receiving the chemotherapy agents while the comparator arm involved all patients receiving the chemotherapy. To allow for the different numbers of patients entering the state transition Markov model, three distinct versions of the model were created for patients receiving:
(i) a positive test result and the precision medicine; (ii) a negative test result and chemotherapy (iii) no test and chemotherapy.

5.3.10 Model Input Parameters

This section describes the key parameters used in the two elements of the decision-analytic model: probabilities for the decision tree; transition probabilities for the Markov model; clinical effectiveness; costs; and consequences. The values used in the baseline model as well as the uncertainty around these estimates are outlined. In addition, any assumptions made about the parameters and their uncertainty are discussed. In general, model input parameters were identified from the published values in the source decision-analytic model-based economic evaluation (234). These values were used to ensure the final estimates of cost-effectiveness were as close to those estimated in the original analysis as possible, thereby ensuring face validity of the case study models. Where model parameters were not available, for the decision tree, information was taken from studies published in academic journals. The model input parameter values for the decision tree are listed in table 5.1 and table 5.2 for the state transition Markov model.

Table 5.1: Model input parameters for decision tree

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value: Base Case</th>
<th>Value: PSA Distribution</th>
<th>Assumptions</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Defining the population</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number of stage III/IV NSCLC patients</td>
<td>16,627</td>
<td>Fixed</td>
<td>Used population of NSCLC patients from above study and removed 15% who are assumed to test positive for <em>EGFR</em> mutations and would therefore receive an <em>EGFR</em> TKI</td>
<td>(234)</td>
</tr>
<tr>
<td>excluding those with <em>EGFR</em> mutations</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>ALK mutation prevalence</strong></td>
<td>3.4%</td>
<td>~Beta(96,2864)</td>
<td>Originally from Bang et al., (2011), a systematic review of 13 original papers. Value for screening all NSCLC patients rather than just adenocarcinoma</td>
<td>(234,245)</td>
</tr>
<tr>
<td><strong>Probabilities</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IHC Sensitivity</td>
<td>95%</td>
<td>~U(90,100)</td>
<td>Weighted average of sensitivity across four testing platforms based on values identified in a systematic review. Used value of 95% in analysis as clinicians believed this value would be reached at optimum utilisation</td>
<td>(128)</td>
</tr>
</tbody>
</table>
IHC Specificity 100% Fixed Weighted average of specificity across four testing platforms based on values identified in a systematic review. Used value of 100% in analysis as clinicians believed this value would be reached at optimum utilisation (128)

Unit costs

<table>
<thead>
<tr>
<th></th>
<th>Base Case Value</th>
<th>Distribution</th>
<th>Assumptions</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>IHC testing</td>
<td>£25</td>
<td>~U(£23, £28)</td>
<td>The cost of testing is redacted in Duarte et al., (2013). A value of £120 was suggested in the qualitative interviews in chapter 5. A recent NICE Medtech innovation briefing suggests that the cost is between £100 and £150 Chapter 5, (246)</td>
<td></td>
</tr>
<tr>
<td>FISH testing</td>
<td>£120</td>
<td>~U(£102,124)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 5.2: Model input parameters for the state transition Markov model

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Base Case Value</th>
<th>Distribution</th>
<th>Assumptions</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Crizotinib Progression Free Survival</strong></td>
<td>Weibull survival curve Shape=1.368 Scale=10.554</td>
<td>Asymptotic normal distribution of survival curve parameter estimates</td>
<td>Based on curve fitted to reproduced KM curve (247)</td>
<td></td>
</tr>
<tr>
<td><strong>Crizotinib Overall Survival</strong></td>
<td>Weibull survival curve Shape=0.952 Scale=34.931</td>
<td>Asymptotic normal distribution of survival curve parameter estimates</td>
<td>Based on curve fitted to reproduced KM curve (247)</td>
<td></td>
</tr>
<tr>
<td><strong>Docetaxel Progression Free Survival</strong></td>
<td>Log-normal survival curve Mean=1.2735 Standard Deviation=1.0091</td>
<td>Asymptotic normal distribution of survival curve parameter estimates</td>
<td>Based on curve fitted to reproduced KM curve (247)</td>
<td></td>
</tr>
<tr>
<td><strong>Docetaxel Overall Survival</strong></td>
<td>Hazard ratio of 0.627 applied to survival curve for crizotinib OS</td>
<td>~N(0.33,1.24)</td>
<td>Actual value redacted in publication. Ad hoc value derived from ratio of median OS for docetaxel to median OS for crizotinib</td>
<td>(234)</td>
</tr>
<tr>
<td><strong>Utilities</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Crizotinib stable disease utility</td>
<td>0.82</td>
<td>1 minus a disutility drawn from the following distribution ~beta(114,520)</td>
<td>Based on “on treatment” utility from original trial</td>
<td>(248)</td>
</tr>
<tr>
<td>Crizotinib progressive disease utility</td>
<td>0.75</td>
<td>1 minus a disutility drawn from the following distribution ~beta(201,601)</td>
<td>Based on midpoint of range from manufacturers submitted sensitivity analysis</td>
<td>(234)</td>
</tr>
<tr>
<td>Chemotherapy stable disease utility</td>
<td>0.74</td>
<td>1 minus a disutility drawn from the following distribution ~beta(216,584)</td>
<td>Based on “on treatment” utility from original trial</td>
<td>(248)</td>
</tr>
<tr>
<td>Chemotherapy progressive disease utility</td>
<td>0.57</td>
<td>1 minus a disutility drawn from the following distribution ~beta(449,595)</td>
<td>Based on midpoint of range from manufacturers submitted sensitivity analysis</td>
<td>(234)</td>
</tr>
<tr>
<td><strong>Costs</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Monthly cost of Crizotinib</td>
<td>£4,689.00</td>
<td>Fixed</td>
<td>Original value from 2012 British National Formulary.</td>
<td>(234)</td>
</tr>
<tr>
<td>Monthly cost of docetaxel</td>
<td>£1,527.86</td>
<td>U~(£1,375, £1,681)</td>
<td>Original value from 2012 British National Formulary.</td>
<td>(234)</td>
</tr>
<tr>
<td>Description</td>
<td>Cost</td>
<td>Confidence Interval</td>
<td>Notes</td>
<td></td>
</tr>
<tr>
<td>-----------------------------------------------------------------------------</td>
<td>----------</td>
<td>---------------------</td>
<td>----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
<td></td>
</tr>
<tr>
<td>Docetaxel administration</td>
<td>£148.20</td>
<td>~U(£131, £160)</td>
<td>Figure based on costs included in NICE STA for pemetrexed. Here it is assumed that there is a constant cost per month rather than different costs in first and subsequent months. In original analysis there was only a difference of £0.81.</td>
<td></td>
</tr>
<tr>
<td>Treatment of adverse neutropenia events in chemotherapy arm</td>
<td>£38.16</td>
<td>~U(£28.62, £47.70)</td>
<td>Based on 4 (range 3-5) days of treatment with granulocyte colony-stimulating factor at £52.71 per day and an incidence of 18.1% in patients receiving chemotherapy.</td>
<td></td>
</tr>
<tr>
<td>Routine medical management cost in stable disease state</td>
<td>£241.44</td>
<td>~U(£217, £266)</td>
<td>Based on values used in previous NICE TA’s.</td>
<td></td>
</tr>
<tr>
<td>Routine medical management cost in progressive disease state</td>
<td>£178.09</td>
<td>~U(£160, £196)</td>
<td>Based on values used in previous NICE TA’s.</td>
<td></td>
</tr>
<tr>
<td>Palliative care before death (one off cost in dead state)</td>
<td>£3,923</td>
<td>Fixed</td>
<td>Based on values used in previous NICE TA’s.</td>
<td></td>
</tr>
</tbody>
</table>

**5.2.10.1 Probabilities**

Values for probabilities were required for decision tree and state transition Markov model. The probabilities required for the decision tree reflect the likelihood that a patient receiving ALK testing receives a positive or negative result. This was determined by the prevalence of ALK mutation in lung cancer patients and the sensitivity and specificity of the IHC and FISH tests. Therefore, IHC and FISH testing parameters were extracted from Djalalov et al (2014) who used values based on a systematic review of 7 studies (128). The authors found a sensitivity of 93% and a specificity of...
99%. The authors expected the sensitivity of the test to be better in practice so used a value of 95%. In this study, the 95% value was used in the base case analysis. FISH testing was assumed to have a sensitivity and specificity of 100% as it was the current gold standard test at the time of submission.

5.2.10.2 Transition Probabilities

The probabilities required for the state transition Markov model reflect the likelihood that a patient would transition between health states in a given cycle. These transition probabilities were derived from the durations of PFS and OS for the intervention and comparator. Parametric survival curves were used to populate the transition probabilities in the treatment Markov models. The PFS curves were used to calculate the probability of moving from the progression-free state to the progressive disease state in each one month interval. The OS curve for crizotinib was used to determine the probability of patients moving from either the progression free or progressive disease states to the dead state in a one month time interval. A hazard ratio was applied to this curve to determine the probability of a patient receiving docetaxel moving from either the progression free or progressive disease states to the dead state. The following section outlines the clinical evidence used to inform the transition probabilities in the Markov model.

5.2.10.3 Clinical Effectiveness

Populating the state transition Markov model in terms of clinical effectiveness data was a challenge because of redacted information in the technology appraisal report and the absence of reported scale and shape parameters for the parametric survival curves used to extrapolate progression free survival and overall survival. To address this problem, an approach proposed by Guyot et al. (2012) was used to reconstruct the survival data for crizotinib and docetaxel from published Kaplan-Meier (KM) curves (249). In this approach, the software package DigitizeIt (250) is first used to digitally read a sample of points on a published KM curve to produce a set of underlying data. Data are extracted across various intervals and an additional file is created with details as to these time intervals, the row numbers of the data in the extracted data pertaining to these intervals and the numbers at risk in each period. The algorithm outlined by Guyot et al. (2012) was then applied to these files to incorporate censoring of the data that was assumed to have occurred at a constant rate. The algorithm was applied in the software R (224) by adapting the authors’
published code and involved the use of the “survival” and “flexsurv” packages (251,252). The output of the algorithm is the approximated individual patient data that would have produced the published KM curve.

As KM curves were not published in the ERG group report for crizotinib, they were taken from the published results of the RCTs used in the submission (PROFILE 1005 AND 1007). Progression free survival data for crizotinib and docetaxel were based on the published results of PROFILE 1007 (253). The overall survival data used in the source model were taken from, the then incomplete, PROFILE 1005 study. As these data were not available, this case study used the mature, final overall survival data from PROFILE 1005 (254).

Using the estimated individual-level patient data from the RCTs, parametric survival curves were fitted to extrapolate the data to longer time-frames. These curves model the probability of PFS or OS at a given time period based on different survival distributions such as the exponential, Weibull, and Gompertz distributions among others (255). The same underlying distribution as assumed by the manufacturer for each curve was used in this case study. The same ranking of appropriate assumed distribution, based on model fit statistics, as reported in the source economic evaluation was obtained in this case study. The fitted parametric survival curves for crizotinib PFS (see figure 5.3), docetaxel PFS (see figure 5.4) and crizotinib OS (see figure 5.5) were generated using the “ggplot2” package in R (256).
Figure 5.3: Weibull survival curve for progression free survival with crizotinib

Footnote: the black solid line indicates the Kaplan-Meier curve created from the reconstructed survival data while the black dotted lines represent the confidence intervals around this curve. The red solid line represents the survival curve estimated from the specified parametric survival curve while the red dotted lines represent the confidence intervals around this curve.
Figure 5.4: Log-normal survival curve for progression free survival with docetaxel

Footnote: the black solid line indicates the Kaplan-Meier curve created from the reconstructed survival data while the black dotted lines represent the confidence intervals around this curve. The red solid line represents the survival curve estimated from the specified parametric survival curve while the red dotted lines represent the confidence intervals around this curve.
Figure 5.5: Weibull survival curve for overall survival with crizotinib

Footnote: the black solid line indicates the Kaplan-Meier curve created from the reconstructed survival data while the black dotted lines represent the confidence intervals around this curve. The red solid line represents the survival curve estimated from the specified parametric survival curve while the red dotted lines represent the confidence intervals around this curve.

To calculate the OS of patients receiving docetaxel, Pfizer applied a hazard ratio (HR) to the parametric survival curve for crizotinib. The hazard ratio measures the rate of survival in one group relative to another at any time point (255). However, the value of this HR was redacted in the published Evidence Review Group report (234). In this case study the HR was approximated by taking the ratio of reported median overall survival in the docetaxel group (20.7 months) to that of those in the crizotinib group (33 months). This resulted in a value of a HR of 0.627.
5.2.10.4 Costs

All prices are reported in 2014 pounds sterling. There were two main types of relevant NHS costs (for the test pathway and for the medicine pathway). A cost of IHC testing of £25 was included in the published NICE TA296 but the cost of FISH testing was redacted. In this case study, a value of £120 was used based on a value stated in a subsequent NICE Technology Appraisal (246) and information supplied during the qualitative interviews in chapter 5.

Cost data for both the intervention and comparator treatments were readily available in the published NICE TA296. The monthly cost of crizotinib was £4,689 while docetaxel was £1,528. As docetaxel is delivered intravenously in hospital it had an administration cost of £148.20 per month. Adverse events were similar across intervention and comparator with the exception of higher rates of neutropenia in those patients receiving chemotherapy. A cost of £38.16 was included for treating this neutropenia based on a calculation included in the Evidence Review Group report of four (range three to five) days of treatment with granulocyte colony-stimulating factor at £52.71 per day and an incidence of 18.1% in patients receiving chemotherapy. Routine medical management costs in the stable and progressive states of both the intervention and comparator were £241.44 and £178.09 respectively. A one-time palliative care cost of £3,923 was applied at time of death for each patient regardless of their treatment modality. This value was identified by calculating the additional number of weeks spent in the dead state in each cycle in each arm to represent the number of patients who had died. This figure was multiplied by the discounted cost of palliative care and added to the total expected cost for the intervention and comparator.

5.2.10.5 Consequences

The relevant consequences used in this case study, consistent with the NICE reference case, was the quality-adjusted life years (QALYs) gained by patients receiving the intervention or comparator. QALYs comprise two components: health-related quality of life (HRQoL), and length of life. The latter was determined by the duration in which patients were in either the progression free or progressed disease health states as opposed to the death health state.

HRQoL is measured using a health state utility value between 1 for full health and 0 for death. Negative values are also possible for health states which are worse than death. The impact on
health was assumed to be a consequence of the medicine not the testing process. Utility values for cancer health states in the treatment Markov models were redacted in the published Evidence Review Group critique. Utility values were instead taken from published results of the PROFILE 1007 study used in the source model (248). On-treatment utility values were used to represent the utility value of the progression free state (crizotinib=0.82, docetaxel=0.74). The utility values used for the progressed disease state in the source model were taken from the average utility of patients on ending treatment in PROFILE 1007. However, these values were not published in the journal article so values for this state were calculated by taking the midpoint value of the lower and upper bound utility values presented in deterministic sensitivity analysis in the source model (234). This corresponded to a value of 0.75 for crizotinib and 0.57 for docetaxel.

5.3.11 Analysis

The analysis was divided into three parts: the base case analysis; sensitivity analysis; and a comparison of the outcomes of the model with those of the model it sought to replicate.

5.3.11.1 Base Case Analysis

In the base case analysis for this study point estimate values for each model input parameter were used to determine the incremental costs and consequences of testing NSCLC patients for NSCLC mutations and then using targeted treatment compared with treating all patients with docetaxel.

5.3.11.2 Discount Rate

Discounting reflects the differential value of costs and consequences which are experienced in the current year as opposed to future years. This differential value arises due to the opportunity cost of spending money or receiving benefits in the current year compared to future years (257). Costs and consequences were discounted at a rate of 3.5% as recommended in the NICE reference case (44).

5.3.11.3 Determining the Incremental Cost-Effectiveness of the Intervention
The primary outcome of the case study models was an estimate of the cost-effectiveness of biomarker testing and targeted treatment compared to no testing and universal chemotherapy. The total costs and benefits of version (i) and version (ii) of the state transition Markov model were summed to calculate the total costs and QALYs of the intervention. These were then compared to the total costs and benefits of the comparator (version (iii)) and the ICER and total net benefit of the intervention was calculated.

An estimate of the cost-effectiveness of these interventions was provided by calculating the incremental cost-effectiveness ratio (ICER) for the intervention (see equation 5.1). This estimate represents the cost per additional quality-adjusted life year gained from the intervention relative to the comparator and is calculated using the formula below:

\[
ICER = \frac{(Cost_2 - Cost_1)}{(QALY_2 - QALY_1)}
\]  
(Equation 5.1)

Where \(C_2\) is the expected total cost of the intervention, \(C_1\) is the expected total cost of the comparator, \(QALY_2\) are the expected total QALYs experienced by those receiving the intervention, and \(QALY_1\) are the expected total QALYs experienced by those receiving the comparator. The estimated ICER is then compared with the cost-effectiveness threshold and if it is lower there is said to be evidence that the intervention is cost-effective and should be introduced. If the ICER is higher than the threshold then the evidence suggests the intervention is not cost-effective and should not be recommended.

In the technology appraisal for crizotinib, the NICE end-of-life threshold of £50,000 was used as the benchmark for cost-effectiveness. This threshold is used for interventions where there is small patient population, with patients who are expected to have less than 2 years left to live, and whose life expectancy is likely to be extended by at least 3 months with use of the intervention (71). This threshold will also be used in the analysis for these case studies

5.3.11.4 Determining Total Net Monetary Benefit of the Intervention

The net monetary benefit of treating all patients with the intervention was calculated to generate an estimate of the total yearly value of the intervention to the NHS produced by the intervention. The
yearly value of the intervention is the expected net monetary benefit per patient multiplied by the number of patients. This calculation provided a baseline estimate for the value of implementation-based approach which will be used in incorporating capacity constraints into the models in the following chapters of this thesis. The estimate produced will be the value of perfect implementation as economic evaluations currently implicitly assume all patients immediately receive the intervention following approval. Equation 5.2 outlines how the net monetary benefit of perfect implementation of the interventions was calculated as outlined in chapter 3:

\[
\text{Value of Perfect Implementation} = \quad n(k. \Delta QALY_{p=1} - \Delta C_{p=1}) - n.p(k. \Delta QALY_p - \Delta C_p)
\]

\[\text{(Equation 5.2)}\]

\[\text{(Where } k = £50,000)\]

Alternatively this can be replace more simply by substituting in NMB which is fixed in the base case analysis:

\[
\text{Value of Perfect Implementation} = \quad n(NMB_{p=1}) - n.p(NMB_p)
\]

\[\text{(Equation 5.3)}\]

or

\[
\text{Value of Perfect Implementation} = \quad n(NMB_{p=1} - p.NMB_p)
\]

As in this example it is assumed that implementation goes from zero to 100% upon approval then this reduces to:

\[
\text{Value of Perfect Implementation} = \quad n(NMB_{p=1})
\]

\[\text{(Equation 5.4)}\]

### 5.3.11.5 Sensitivity analysis

To account for parameter uncertainty affecting the estimated ICERs and value of the interventions, probabilistic sensitivity analysis was undertaken. This process involves assigning distributions to the model parameters and running a specified number of simulations with new parameter values from these distributions drawn in each iteration. The distributions used for each parameter are outlined for the decision tree (Table 5.1) and state transition Markov model (Table 5.2). In this case
study, 1,000 such Monte Carlo simulations were completed. The ICERs estimated in each iteration were used to generate net monetary benefit values. The proportion of simulations with positive net benefit estimates in each iteration was then used to calculate the probability that the intervention would be cost-effective.

Evidence on the assumed distribution of the values of key parameters was rarely available in the published Evidence Review Group report on the submission by Pfizer. The distribution of prevalence rate of ALK mutations was reported based on a published review (245). The sensitivity of IHC testing reported in Djalalov et al., (2014) featured upper and lower bounds but no information about the distribution so a uniform distribution was applied to these bounds of the parameter estimate. The recreation of the survival curves from the published KM curves meant that estimates of the survival distributions could be created. As the hazard ratio for overall survival for docetaxel was produced using an approximation, a true distribution could not be estimated. Instead, this analysis used the reported lower and upper bounds for deterministic sensitivity analysis conducted by the manufacturer to represent confidence intervals for this variable. From these, a normal distribution of hazard ratios was created as an approximation for the distribution of hazard ratios. This approach was also used for the cost parameters, although uniform distributions were used to provide conservative over-estimates of the potential uncertainty. For the utility estimates for the progression free disease states, confidence intervals were available from the published results of the PROFILE 1007. Beta distributions were fitted using the mean utility and these confidence intervals. For the approximated utility values used for the progressive disease state, it was assumed that the confidence intervals would be of the same magnitude and these were calculated and used to create additional beta distributions for these parameters.

5.3.11.6 Model Validation

The evidence provided by decision-analytic model based economic evaluations informs decisions as to the allocation of significant levels of health system resources with resulting changes in the health of different groups of patients. As such it is important to evaluate the validity of decision-analytic models. To determine the validity of the case study model, the Assessment of the Validation Status of Health-Economic decision models (ADVISHE) model validation checklist was applied (258). This checklist poses a series of questions to modellers about various aspects of model validity including in the conceptual model, parameters used and the cross validity of the model and results. However, as the aim of this model was to replicate the manufacturer submission
by Pfizer for the NICE Technology Appraisal of crizotinib (NICE TA 296) (234), some aspects of validity will not be relevant in this case study while others will be of greater importance. For example, it is assumed that the manufacturers would have addressed some components of model validation: validation of the conceptual model (Part A of the AdViSHE checklist) and input data validation in the source model (Part B of the AdViSHE checklist). In this case study, model validation therefore focused: validation of the computerised (case study) model (Part C of the AdViSHE checklist) and operational validation (Part D of the AdViSHE checklist). The key purpose of the model validation conducted in this study was to ensure the results of the baseline model were similar to the results produced in the source model. The outputs on which the validity of the model were judged were the expected costs and QALYs of the intervention and comparator per patient treated, the incremental costs and benefits, and the ICERs.

An additional potential approach to improving the validity of the replicated model was to calibrate the model to the outputs of the source model (259). This involves changing the input parameters, particularly those that are unknown or uncertain, in order to generate outputs that align with some validated data. In this case, unknown input parameters such as the health utility values or parameters of the survival curves could be changed in the case study model in order to attempt to generate the outputs of the source model. However, as the outputs to be matched in this study were the expected costs and QALYs of the intervention and comparator, no calibration was undertaken. This was because QALYs are a composite measure and their expected value was driven by two key unknown parameters: the health utility for the Markov model health states; and the parameters of parametric survival curves. As such there may be different combinations of these parameters which yield QALY values in the case study model which match those produced in the source model and there would be no way to test which combination was correct meaning it was not possible to use any of the proposed methods of calibration (260).

5.4 Results

This section presents the base case results; results of the sensitivity analysis; and a comparison of the outcomes of the model with those of the model it sought to replicate.

5.4.1 Base Case Analysis
In the base case analysis, the use of ALK testing to guide treatment with crizotinib or docetaxel had an incremental cost of £1,391 (using a price year of 2014) and provided an additional 0.035 QALYs per patient tested. When framed in terms of only ALK positive patients, as in the source model, crizotinib provided an additional 1.044 QALYs at an incremental cost of £40,161. Depending on this framing, the ICER for the intervention was £39,198 or £38,468 respectively. The higher ICER when accounting for the testing process results from the cost of testing and imperfect nature of the test. For example, 16,032 patients received negative results from the original IHC test whilst 30 received initial positive results which were later determined to be false-positive by FISH tests. Such patients incur a cost of testing but receive no expected benefit so they serve to inflate the ICER.

This case study used the example of ALK testing to inform prescribing of crizotinib and the ICER of £39,198 per QALY represents the cost-effectiveness of the intervention. Given this ICER, there is evidence to suggest that ALK testing and treatment with crizotinib or docetaxel is cost-effective assuming the NICE end-of-life threshold of £50,000 per QALY. At this level of cost-effectiveness, the intervention would provide a value (total NMB) of £6,373,887 per year. Figure 5.6 shows the value of the intervention at different cost-effectiveness thresholds.

**Figure 5.6: Total incremental monetary net benefit of ALK testing to guide treatment with crizotinib or docetaxel by threshold level**
5.4.2 Sensitivity Analysis

The probabilistic sensitivity analysis for this case study suggested that there was an 85% probability that ALK mutation testing and targeted treatment with crizotinib was cost-effective assuming the NICE end-of-life threshold. Figure 5.7 shows the cost-effectiveness acceptability plane with the 1,000 simulated incremental costs and benefits. Figure 5.8 shows the distribution of total annual value estimates for the intervention.

Figure 5.7: Distribution of Estimated Incremental Costs and Benefits
Figure 5.8: Distribution of Annual Total Expected Net Monetary Benefit

5.4.3 Model Validity

To determine the validity of the case study model, Part C and D of the ADVISHE checklist were used to critique the model and the estimated results (258). Part C of the checklist concerns validation of the technical production of computerised model to check for coding errors. This was completed by a second researcher.

To further check for coding errors, the model was run with various extreme parameter values including large ranges in drug cost, utility values for stable and progressive disease states, and differing parametric survival curves. The component decision tree and state transition Markov model were individually checked for errors. The conditional probabilities for the decision tree were
calculated and manually checked. These conditional probabilities were applied to the cohort to determine patient flows into the different Markov models and these numbers exhibited validity when compared to the estimated number of patients receiving crizotinib in the source model. The Markov traces for each treatment option were examined and plotted over time for evidence of coding errors.

Part D of the ADVISHE checklist focuses on operational validity. The key area of validity for the case study model is in cross validation testing of model outcomes when compared with the source model. For the case study model to be valid, it should estimate similar estimates of deterministic expected cost and consequences to those produced by the source models for patients receiving either the intervention or comparator. As the decision problem used in the case study differed slightly to that in the source model, the expected cost and QALYs for patients who had received testing and were ALK positive and received crizotinib were compared with those patients who had tested negative and received docetaxel (see table 5.3).

Table 5.3: Comparison of the deterministic results from this case study with published manufacturer estimates

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Source model (Percentage difference)</th>
<th>Case Study Model (Percentage difference)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean Total Cost for Patients Receiving Docetaxel</td>
<td>£13,922</td>
<td>£14,748 (+6%)</td>
</tr>
<tr>
<td>Mean Total Cost for Patients Receiving Crizotinib</td>
<td>£54,149</td>
<td>£54,908 (+1%)</td>
</tr>
<tr>
<td>Mean Total QALYs for Patients Receiving Docetaxel</td>
<td>0.981</td>
<td>1.049 (+7%)</td>
</tr>
<tr>
<td>Mean Total QALYs for Patients Receiving Crizotinib</td>
<td>1.949</td>
<td>2.093 (+7%)</td>
</tr>
<tr>
<td>Incremental Costs</td>
<td>£40,227</td>
<td>£40,161 (-0.2%)</td>
</tr>
<tr>
<td>Incremental QALYs</td>
<td>0.968</td>
<td>1.044 (+8%)</td>
</tr>
<tr>
<td>ICER (£/QALY)</td>
<td>£41,554</td>
<td>£38,468 (-7%)</td>
</tr>
</tbody>
</table>

The expected costs and QALYs produced in this case study slightly underestimated the ICER produced in the source model. This effect appears to mainly be driven by larger estimated QALYs and an increase in the difference between these estimates. One explanation for this effect may be that the mid-point values taken from the published confidence intervals which were used to
estimate some utility values in the case study model were overestimates of the deterministic health utility. It may be that the distribution of HRQoL for patients with NSCLC is skewed to the left, meaning that patients are more likely to have more severe symptoms, with fewer patients experiencing mild symptoms. This would have meant that the true average health utility for some of the health states may have been lower, reducing the total QALYs experienced in both arms.

The probability of cost-effectiveness, estimated in the probabilistic sensitivity analysis, was greater than the 65% estimated in the manufacture submission by Pfizer but this is expected given the lower ICER estimated in the case study. Additional differences are likely driven by the distribution used for the hazard ratio of overall survival for docetaxel relative to crizotinib. This value was only an approximation in the deterministic analysis and was shown to be the critical variable in the source model. In this analysis the distribution of the ratio was not available so a normal distribution was chosen for pragmatic reasons. However, other distributions will lead to very different probabilities of cost-effectiveness. For example, using a uniform distribution bound by the intervals reported in the deterministic sensitivity analysis in the source economic evaluation yields a probability of cost-effectiveness of only 49%.

5.5 Discussion

This case study detailed the construction of a replica decision-analytic model to evaluate the cost-effectiveness of an example of precision medicine, using ALK testing and crizotinib, in NSCLC. The decision-analytic model produced was a sufficiently good replication of the decision-analytic model-based cost-effectiveness analysis produced by the manufacturers of crizotinib in their submission to NICE and provided estimates of the costs and consequences of the intervention and comparator. The case study provided a relatively accurate prediction of the incremental costs and QALYs of the intervention when compared to the source model. The estimated costs of the intervention and comparator differed by 1% and 6% respectively while the estimated consequences of both strategies were 7% higher than the original model. Similarly, the ICER estimated for ALK testing and targeted treatment differed from that in the manufacturer submission by 7%. The observed differences in expected costs and QALYs (range of 1 to 7%) was deemed to be minimal because ALK testing and crizotinib was estimated to be cost-effective assuming the NICE end-of-life threshold.
This result was expected given that the analysis was based on a reproduction of that included in application by Pfizer for approval of crizotinib by NICE. In the NICE appraisal of crizotinib, the decision-analytic model was later critiqued by an external Evidence Review Group who produced a range of much higher estimated ICERs. At the time of the first submission to NICE, crizotinib was not recommended due to the uncertainty regarding the hazard ratio for the overall survival of patients receiving docetaxel. However, following a re-analysis by the Cancer Drugs Fund in 2016 using mature survival data, updated unit costs, and a revised patient access scheme, crizotinib was approved for patients as a second line treatment (261).

A recent study by McManus et al., (2019) explored the difficulty in reproducing published economic evaluations (262). In the study, the authors attempted to replicate five examples of published economic evaluations with mixed results. The estimates of incremental costs produced by the authors varied by -4.54% to 108.00% while the estimated benefits varied by -3.81% to 0.50%. The observed differences between the estimated benefits in this case study and the source model were likely driven by difficulties in reproducing the effectiveness data and utility values in the ALK model due to their redaction in the ERG report. The tendency for all outcomes to be overestimated suggests that the reproduced survival curves used in this case study may have slightly overestimated survival for patients receiving both the intervention and comparator. The lower estimated ICER resulting from this case study meant that the probability of cost-effectiveness of the intervention was higher than in the source model (85% vs 65%). This effect was also partly driven by the approximation of the hazard ratio for overall survival for docetaxel and the distribution of this parameter.

The findings of this case study and the source model in terms of the cost-effectiveness of ALK testing and crizotinib may be misleading due to the omission of health system capacity constraints. The analysis assumed that all patients received testing and that there were no fixed costs in setting up ALK testing in the NHS. These factors were omitted from the analysis submitted to NICE even though “the committee heard from the clinical experts that there (was) currently no established ALK testing strategy in UK clinical practice” (261). In addition, the quality of the test and testing process in clinical practice may have differed with that reported in the original trials used to provide evidence of the clinical effectiveness of crizotinib (247,254). This difference in test performance may have led to poor stratification of the treatment with reduced expected outcomes for patients. The estimated cost-effectiveness and annual value of ALK testing to target treatment with crizotinib or docetaxel may have been optimistic estimates.
5.6 Conclusion

This case study outlined the construction of a replica decision-analytic model-based cost-effectiveness analysis of ALK testing and crizotinib for patients with NSCLC. The analysis performed well with regards to predicting similar incremental costs and QALYS to those produced the source model produced to inform NICE TA296. This was achieved even in the absence of a fully executable model and a substantial amount of redacted information in the original submission. The case study model suggested that ALK mutation testing and targeted treatment with crizotinib would be cost-effective in the absence of health system capacity constraints.

5.7 Summary of Chapter 5

Chapter 5 reported a case study that created a decision-analytic model-based cost-effectiveness analysis of ALK testing to guide treatment with crizotinib. This model produced baseline estimates of the cost-effectiveness and total net benefit of the intervention in the absence of capacity constraints. The results obtained from the model were similar to those in the original model it sought to replicate.

In chapter 6, relevant capacity constraints which were present during the introduction of ALK testing will be selected based on the findings of chapter 4. The model reported in this study will then be adapted to incorporate these capacity constraints. The impact of these constraints on the ICER and total net benefit of ALK testing will be determined using that static value of implementation method developed in chapter 3 and these results will be compared to those generated by the baseline model reported in this chapter.
Chapter 6

Incorporating Capacity Constraints into a Decision-Analytic Model-Based Cost-Effectiveness Analysis of ALK Testing to Guide Treatment with Crizotinib in NSCLC

6.1 Chapter 6 Summary

Chapter 6 reports a study which builds on the baseline decision-analytic model-based CEA of using ALK testing to guide treatment with crizotinib in NSCLC (see Chapter 5) to quantify the impact of selected capacity constraints (informed by Chapter 4). The base line decision-analytic model was restructured to incorporate the capacity constraints, and the additional parameters required were identified. The incremental costs and QALYs were then re-estimated to determine the impact of the capacity constraints on the relative cost-effectiveness and value (net monetary benefit) of using ALK testing to guide treatment with crizotinib for people with NSCLC.

6.2 Introduction

The use of crizotinib guided by ALK testing is recommended for patients with NSCLC based on NICE guidance (261). However, this technology appraisal of crizotinib did not account for capacity constraints to the introduction of ALK testing. In chapter 2, it was identified that the inclusion of capacity constraints in an economic evaluation could affect the estimated cost-effectiveness of the intervention and optimal treatment strategy. In chapter 4, qualitative interviews with stakeholders in the provision of examples of precision medicine in NSCLC identified that there were a number of such capacity constraints to introducing ALK testing to guide crizotinib in the NHS. This chapter reports a study which sought to determine the impact of incorporating key capacity constraints to the introduction of ALK testing into a decision analytic model which evaluated the cost-effectiveness of ALK testing and targeted crizotinib.

6.2.1 Aims and Objectives

The aim of this study was to quantify the impact of selected capacity constraints on the relative cost-effectiveness of an example of precision medicine in NSCLC.
The objectives of this study were to:

- select relevant capacity constraints to quantify the potential impact on the relative cost-effectiveness and value (net monetary benefit) of ALK testing to target crizotinib;
- restructure an existing baseline (see Chapter 5) decision-analytic model-based cost-effectiveness analysis of ALK testing and crizotinib;
- identify relevant data to re-parameterise the restructured decision-analytic model-based cost-effectiveness analysis of ALK testing and crizotinib to incorporate the selected capacity constraints;
- compare the estimated cost-effectiveness and net monetary benefit produced by the decision-analytic model-based cost-effectiveness analysis of ALK testing and crizotinib incorporating constraints to those produced by the baseline model.

6.3 Methods

This section describes how the baseline decision-analytic model-based cost-effectiveness analysis of ALK testing and crizotinib (see Chapter 5) was adapted to incorporate capacity constraints. In order to incorporate the three capacity constraints, four new decision problems were specified (see section 6.2.2) to explore the cost-effectiveness of the intervention in the presence of each constraint separately and then in combination. This study estimated the static impact of the capacity constraints (see also Chapters 2 and 3 for a definition of static impact). Estimating the static impact means that the study explores the impact of the capacity constraints on the cost-effectiveness of the intervention at the time of approval for a single year’s cohort of patients. Unless otherwise specified in this chapter, the key characteristics of the model such as the study perspective (NHS), time horizon (life-time), price year (2014) and discount rate (3.5%) are the same as those reported in the baseline model in chapter 5.

6.3.1 Selecting Relevant Capacity Constraints

The five themes identified in the qualitative interview study (see Chapter 4) were used as the starting point to select relevant capacity constraints to be included in this quantitative analysis of the impact on relative cost-effectiveness of an example of precision medicine. The capacity constraints were chosen based on discussion in the research team about which constraints were most likely to have a significant impact on the cost-effectiveness of net monetary benefit of the
intervention and based on the ability of the NHS to take actions to address these constraints. This meant that some constraints such as a lack of available tissue sample for conducting mutation testing and delays due to sequential testing of samples were not included. The availability of tissue may to some degree depend on the skill of the physicians taking the tumour sample but is likely to depend more on the health of the patient, and at the time of assessment of crizotinib, techniques such as panel testing and next generation testing which could have removed the problem of sequential testing were not widely available in the NHS.

Three key capacity constraints were selected for this analysis.

i)  *Lack of awareness about how ALK testing was commissioned.* Participants in the qualitative interviews (see Chapter 4) suggested that there was a lack of understanding about how ALK testing was supposed to be reimbursed. This lack of clarity resulted in some instances when no testing was conducted and some instances when testing was conducted but a financial loss may have been experienced by the hospital trust providing the test. These two potential actions may result in different potential economic impacts. If laboratories did not offer testing so as not to experience a financial loss, then a number of patients who may have benefitted from targeted treatment may have received the less effective chemotherapy as they could not receive testing. If hospital trusts offered the testing at a financial loss, then patients would have received testing but other patients receiving interventions from the same budget area may have experienced a health loss from a lack of available funding for other existing interventions. In this scenario the former approach, whereby it is assumed that patients are not provided with testing due to the constraint, was assumed to take place.

ii)  *Degree of centralisation of immunohistochemistry testing.* When the intervention was initially introduced some ALK testing was performed in larger centralised laboratories whilst some was performed in smaller, local laboratories. Due to the volume of requests for ALK testing received at centralised laboratories, turnaround times were potentially slow. It was believed by some interviewees (see Chapter 4) that localising IHC testing to the hospitals where patients are based could reduce turnaround times, ensuring there is less delay in patients starting treatment. Other interviewees (see Chapter 4) believed that centralised testing resulted in economies of scale,
reducing the cost per patient. Some interviewees believed that centralised testing meant that pathologists gained more experience in conducting tests and that this might lead to a better quality of testing than if testing was localised. It is not clear what the impact of using localised or centralised testing would have on the relative cost-effectiveness of ALK testing for crizotinib.

iii) Ability to conduct ALK testing dependent on whether centralised or localised services were used. Interviewees suggested (see Chapter 4) that a lack of financial and human resources meant that it was difficult to offer responsive, high quality testing. The actual impact of limited pathology laboratory capacity is difficult to enumerate. One assumption is that the negative impacts of either centralising or localising pathology testing occur as a result of the limited ability, due to financial and human resource constraints, of pathology laboratories to deliver responsive and high quality testing. Specifically, the longer turnaround times faced by centralised laboratories and the more expensive, lower quality tests associated with localised testing are a product of the capacity of these laboratories.

6.3.2 Defining the New Decision Problems

To incorporate the three selected (see section 6.2.1) capacity constraints into an existing decision-analytic model-based cost-effectiveness analysis of ALK testing and targeted treatment, four new decision problems were specified:

Decision problem 1: What is the cost-effectiveness and net monetary benefit of ALK testing to guide treatment with crizotinib given that some pathology laboratories are unaware of how ALK testing is commissioned?

Decision problem 2: What is the cost-effectiveness and net monetary benefit of ALK testing to guide treatment with crizotinib given that some ALK IHC testing is conducted in centralised pathology laboratories and some is conducted in localised pathology laboratories?

Decision problem 3: What is the cost-effectiveness and net monetary benefit of ALK testing to guide treatment with crizotinib given that some ALK IHC testing is conducted in centralised
pathology laboratories and some is conducted in localised pathology laboratories, and that all types of pathology laboratories are under-resourced?

Decision problem 4: What is the cost-effectiveness and net monetary benefit of $ALK$ testing to guide treatment with crizotinib given that some pathology laboratories are unaware of how $ALK$ testing is commissioned, some $ALK$ IHC testing is conducted in centralised pathology laboratories and some is conducted in localised pathology laboratories, and that all types of pathology laboratories are under-resourced?

### 6.3.3 Restructuring the Decision-Analytic Model

To incorporate the three selected (see section 6.2.1) capacity constraints into an existing decision-analytic model-based cost-effectiveness analysis of $ALK$ testing and targeted treatment, the basic structure of decision tree was altered to be able to address the four decision problems associated with including each capacity constraint individually or in tandem.

**Decision problem 1: What is the cost-effectiveness and net monetary benefit of $ALK$ testing to guide treatment with crizotinib given that some pathology laboratories are unaware of how $ALK$ testing is commissioned?**

The possibility for hospital trusts not to offer IHC testing due to a lack of awareness of commissioning guidelines meant that in the intervention arm, some patients received testing and therefore targeted treatment, while others who were based at trusts who did not know how to be reimbursed for testing could only receive universal docetaxel. In the decision tree outlined in section 5.3.7 (see figure 5.1), incorporating the lack of commissioning awareness involves adding an additional branch for patients who receive chemotherapy because their pathology lab does not offer testing as it is not clear how they would be reimbursed (figure 6.1).

**Decision problem 2: What is the cost-effectiveness and net monetary benefit of $ALK$ testing to guide treatment with crizotinib given that some $ALK$ IHC testing is conducted in centralised pathology laboratories and some is conducted in localised pathology laboratories?**

In the baseline decision-analytic model (see Chapter 5), it was implicitly assumed that all tests were returned with a short turnaround time and that testing was of a high quality. In the qualitative
interviews in Chapter 4, it was identified that either centralising or localising tests could have consequences for the turnaround time or quality of testing. It was suggested that conducting testing in localised laboratories could result in faster turnaround times but testing may be of lower quality. Testing may also be more expensive as smaller laboratories do not benefit from economies of scale. Centralising testing may avoid the issues of higher test cost and lower quality but turnaround time for tests may be longer. In order to incorporate the capacity constraints of slow test turnaround due to centralisation or the lower quality testing associated with localisation, additional branches were added following the conducting of the test to represent the different geographical set-up of the laboratories. Following the receipt of test results from either type of laboratory, the patient experience was the same as in the baseline model.

**Decision problem 3: What is the cost-effectiveness and net monetary benefit of ALK testing to guide treatment with crizotinib given that some ALK IHC testing is conducted in centralised pathology laboratories and some is conducted in localised pathology laboratories, and that all types of pathology laboratories are under-resourced?**

A report in 2016 by Cancer Research UK provided extensive detail about the issue of limited pathology capacity in the NHS (213). However, no evidence was provided to quantify the potential impact on services. Similarly, a paper investigating the potential efficiency improvements in pathology laboratories only highlighted potential cost savings and not gains in outcomes (216). In this study it was assumed that the negative effects of centralised or localised ALK testing occurred because of the limited capacity of the pathology labs. Including the impact of the limited pathology capacity constraint is not visible in the structure of the decision tree as it only acts on the costs and consequences for patients rather than the model structure.

**Decision problem 4: What is the cost-effectiveness and net monetary benefit of ALK testing to guide treatment with crizotinib given that some pathology laboratories are unaware of how ALK testing is commissioned, some ALK IHC testing is conducted in centralised pathology laboratories and some is conducted in localised pathology laboratories, and that all types of pathology laboratories are under-resourced?**

The restructured decision tree including all of the three capacity constraints is shown in figure 6.1. When compared with the decision tree included in the baseline model (see figure 5.1) there are additional branches which allow for some patients to not receive testing because the pathology
laboratory does not know how to commission testing and for patients to receive testing from localised or centralised pathology laboratories. As in decision problem 3, no restructuring is required to incorporate the pathology laboratory capacity as this constraint only acts in changing the parameters in the model.

**Figure 6.1: Decision Tree to Incorporate Capacity Constraints**

The resulting restructured model (figure 6.1) which incorporates capacity constraints appears to be significantly different to that used in the baseline model. However it should be noted that if all patients have access to commissioned IHC testing the decision problem becomes the same as in the baseline model. If there is no limit to pathology capacity and testing is centralised then the model fully collapses to the baseline model. Testing must be centralised for this to be the case as it is assumed that IHC testing is more expensive.
6.3.4 Parameterising the Impact of the Three Capacity Constraints

In addition to altering the structure of the decision-analytic model, the inclusion of capacity constraints meant that additional parameter values needed to be identified and incorporated. Each constraint had the potential to influence the probability of patients experiencing different events of health states in the model, the costs of testing and treating patients, and the outcomes they experienced. In order to identify potential values for these new parameter values, pragmatic, targeted searches were conducted in the Embase and MEDLINE databases (in 2019) using the Ovid search tool (263). Key terms included in the search included “ALK”, “anaplastic lymphoma kinase”, “provision”, “availability”, “access”, “capacity”, “barrier” and “patholog*”. To limit the number of papers to be screened, key words were required to appear in the titles of included papers. Grey literature reports identified previously in this thesis, such as the National Lung Cancer Audit and Cancer Research UK reports, were also used to identify potential parameter values (99,213,264,265). The following section outlines the parameters used in the model to represent the impact of the capacity constraints.

6.3.4.1 Lack of Awareness About how ALK Testing was Commissioned

For the ‘commissioning awareness’ capacity constraint, an estimate of the proportion of patients being treated at trusts where there was knowledge of test commissioning was required in order to produce a probability that a patient would be treated at a trust with knowledge of test commissioning. In the absence of available data on the degree of knowledge of commissioning awareness, the proportion of eligible patients receiving ALK testing in 2014 was used as a proxy. This approach made the implicit assumption that a lack of awareness of test commissioning was the only reason patients did not receive ALK testing.

It was known that UK based data are “lacking on the uptake of tests, their turnaround times, their impact in changing patient management and responses” (265). A pragmatic search of the MEDLINE and EMBASE databases (in 2019) using Ovid using search terms such as “ALK OR anaplastic lymphoma kinase” and “uptake” or “provision” identified a single relevant paper which had examined the availability of ALK testing in nine countries between 2011 and 2013 (266). The proportion of patients receiving ALK testing ranged from 2% (Taiwan) to 37% (Switzerland). For this case study a weighted average of these uptake estimates (23%) was taken to provide an estimate of the availability of ALK testing in the UK. One minus this value (77%) was included as
the probability of patients receiving no testing and chemotherapy rather than ALK IHC and FISH testing and targeted treatment. It was assumed that the commissioning awareness capacity constraint had no other effects on the model probabilities, costs, or QALYs.

6.3.4.2 Degree of Centralisation of Immunohistochemistry Testing

As was illustrated in figure 6.1, the inclusion of the ‘localisation or centralisation’ constraint meant adding separate branches for each location of testing. The probability of patients receiving testing at a hospital which offered testing in-house (localised testing) or through another laboratory (centralised testing) was informed by the National Lung Cancer Audit. This audit is a survey of outcomes for lung cancer patients in all health trusts in the UK (267). In the National Lung Cancer Audit conducted in 2014, it was reported that 27% of trusts offered ALK testing in-house (268). As there were 176 trusts at the time of the report, this suggests that 48 trusts had in-house ALK testing in 2014. Under a policy of centralisation there will be a reduced, optimum, number of pathology labs offering testing. Given the recent move to consolidate pathology laboratories into 29 networks in the UK, it was assumed that one laboratory in each network would offer ALK testing in an ideal scenario (214). Therefore it is assumed that under full centralisation, ALK testing would be offered by 29 laboratories and under full localisation testing would be offered in 176 laboratories. In this example it is assumed that patients are distributed evenly between laboratories meaning that 11% (19 out of 176) of patients would receive localised testing and 89% would receive centralised testing in the base case.

In the qualitative interview study reported in chapter 4, some participants believed that testing would be cheaper if it was performed in centralised laboratories. A pragmatic search (conducted in 2019) using terms including “pathology*”, “capacity” and “barrier” identified a study of the cost of pathology laboratories suggested that centralising testing could reduce pathology costs by up to 17% (216). To reflect this, the cost of IHC testing for patients receiving localised testing was inflated by 17% to a value of £29.25.

The potential negative aspect of localised testing was identified by participants in the qualitative interview study as potentially poorer quality ALK testing. The potential impact on QALYs of reduced test quality in localised laboratories was modelled using data from the PROFILE 1005 study which was included as a key source of evidence in the NICE technology appraisal of crizotinib (227,234,254). In the trial providing evidence of clinical effectiveness of crizotinib, ALK
testing to determine patient eligibility was originally conducted in central laboratories before later being rolled out to localised laboratories. Patients in the local \textit{ALK} testing group had lower median progression free survival (PFS) (6.9 versus 8.4 months) and overall survival (OS) (16.9 versus 21.8 months) than those in the centralised group. It is assumed in this study that the differences in PFS and OS between patients whose tests were processed in localised or centralised laboratories is solely caused by the quality of the testing process in these laboratories. However, in reality there may be a number of confounding factors for these estimates such as the demographic characteristics of the groups receiving localised or central testing and the quality of treatment available in these settings.

To incorporate these effects, the transition probabilities for the state transition Markov model associated with patients receiving localised testing were altered. Simple hazard ratios were calculated for PFS (0.822) and OS (0.775) based on the ratio of the median survival in months. These estimates were applied to the transition probabilities between the progression free, progressive disease, and dead health states used in the state transition Markov model (see figure 5.2 in Chap 5).

While centralised laboratories were perceived by interview participants (see section 4.3.4) to provide better quality tests, concerns were raised about their ability to provide a rapid turnaround time for test results. This delay may mean patients have to start their cancer treatment later as they are waiting for test results. Using a rapid review (in 2019) and search terms for “lung cancer”, “treatment”, and “delay” yielded a number of papers which have sought to quantify the impact of delays to the start of cancer treatment on treatment outcomes (269–271). These studies found an unexpected result in the correlation of longer delays to treatment start with better outcomes for patients. This effect can be explained by the tendency to prioritise the start of treatment for patients with more severe illness (270). As these patients are likely to have a shorter life expectancy, a correlation was found between shorter times to starting treatment and poorer health outcomes.

To determine whether delayed treatment start may have an impact on patient outcomes having controlled for disease severity, two consultant oncologists were contacted. The oncologists believed that receiving IHC results in two weeks (the delay duration suggested in the interview study) rather than three days would not have an impact on treatment outcomes for patients. However, they did believe that the delay would result in patients having to have an additional appointment with their consultant to receive results and the delay may also cause anxiety for
patients which could be reflected in a temporarily reduced quality of life. A cost of £101 was assigned to patients receiving centralised ALK testing based on the cost of an assumed additional hour long consultant oncologist appointment (272).

A pragmatic search (in 2019) using the search terms for “utility”, “cancer”, “delay”, and “diagnosis” identified a paper which suggested that patients’ HRQoL increases by 0.03 on receiving a definitive cancer diagnosis (273). It was assumed that the impact of delaying the start of treatment would have the same size impact as waiting for a cancer diagnosis. A disutility of 0.03 was therefore applied to the health related quality of life (HRQoL) for patients receiving centralised testing to represent the anxiety of having to delay their treatment start. The disutility was applied for nine days which is the difference between a three day and two week turnaround time as suggested in the interview study.

6.3.4.3 Ability to Conduct ALK Testing Dependent on Whether Centralised or Localised Services were Used

In this study it was assumed that the negative effects of centralised or localised ALK testing occurred because of the limited capacity of the pathology laboratories. In other words, if there was sufficient capacity in centralised laboratories then tests could be turned around quickly enough to mean that patients would not experience anxiety from delayed treatment start and would not have an additional visit with a consultant. If testing was localised then fully resourced pathology laboratories would be able to conduct testing of equal quality to that of centralised laboratories.
Table 6.1: Parameter Values for the Identified Capacity Constraints

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Constraint</th>
<th>Base Case Value</th>
<th>Distribution</th>
<th>Sources</th>
<th>Assumptions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Probabilities</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Proportion of patients at trusts where there is an awareness of test commissioning arrangement</td>
<td>1) Lack of awareness about how ALK testing was commissioned</td>
<td>0.23</td>
<td>Beta~(498,1660)</td>
<td>Ess et al., (2017) (274) and Lee et al (2018) (266)</td>
<td>Synthesis of ALK testing rates between 2008 and 2013 for 9 countries</td>
</tr>
<tr>
<td>Proportion of patients at trusts with access to localised testing</td>
<td>2) Degree of centralisation of immunohistochemistry testing</td>
<td>0.11</td>
<td>Beta~(19,158)</td>
<td>National Lung Cancer Audit 2014 (268) and NHS Improvement 2019 (214)</td>
<td>The 2014 NLCA suggested that 27% (48) of trusts offered in-house testing for ALK. However, this figure cannot be used as the parameter</td>
</tr>
</tbody>
</table>
because under a centralised test provision there will be a correct number of central pathology labs which offer the test which will correspond to a localisation rate of 0%. A recent strategy by NHS improvement seeks to consolidate the UK pathology laboratories into 29 networks. It is assumed in this study that each network
will have one laboratory which offers *ALK* testing. Therefore in this study it was estimated that 19 trusts offered testing which were not specialised in *ALK* testing. If patients are distributed evenly across trusts then this means 11% would receive localised testing.

<p>| Hazard Ratio of Local Crizotinib | 2) Degree of centralisation of immunohistochemistry testing | 0.82 | Uniform~(0.697,1) | Blackhall et al., 2017 (254) |</p>
<table>
<thead>
<tr>
<th>PFS to Central Crizotinib PFS</th>
<th>3) Ability to conduct ( ALK ) testing dependent on whether centralised or localised services were used</th>
<th>Hazard Ratio of Local Crizotinib OS to Central Crizotinib OS</th>
<th>0.77</th>
<th>Uniform((0.633,1))</th>
<th>Blackhall et al., 2017 (254)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Utilities</td>
<td>2) Degree of centralisation of immunohistochemistry testing [3) Ability to conduct ( ALK ) testing dependent on whether centralised or localised services were used</td>
<td>0.03</td>
<td>Triangular((0, 0.1, 0.03))</td>
<td>Moseholm et al (2016) (273)</td>
<td>Original value from HRQoL gain after confirmed cancer diagnosis. Assumed similar anxiety is experience while</td>
</tr>
<tr>
<td>Costs</td>
<td>2) Degree of centralisation of immunohistochemistry testing</td>
<td>£101</td>
<td>Fixed</td>
<td>PSSRU Unit Costs of Health and Social Care (2016) (275)</td>
<td></td>
</tr>
<tr>
<td>----------------------------------------------------------------------</td>
<td>-------------------------------------------------------------</td>
<td>------</td>
<td>-------</td>
<td>--------------------------------------------------------</td>
<td></td>
</tr>
<tr>
<td>Cost of an extra appointment with an oncologist due to test delay</td>
<td>3) Ability to conduct ALK testing dependent on whether centralised or localised services were used</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cost of IHC testing when conducted in a local lab</td>
<td>2) Degree of centralisation of immunohistochemistry testing</td>
<td>£29</td>
<td>25x(1+uniform~(0.1,0.25))</td>
<td>Buckell et al (2015) (216)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>3) Ability to conduct ALK testing dependent on whether centralised or localised services were used</td>
<td></td>
<td>Original study does not present information from which value could be derived</td>
<td>Cost is 17% higher than that in centralised labs due to inefficiency</td>
<td></td>
</tr>
</tbody>
</table>

### 6.3.5 Analysis

This section reports the methods used for a base case analysis, a sensitivity analysis and model validation.
6.3.5.1 Base Case Analysis

The base case analysis incorporating capacity constraints took part in two stages. The first stage of the analysis sought to determine the impact of each capacity constraint individually while the other two constraints were not present. The second stage of analysis aimed to quantify the combined impact of all three capacity constraints at the same time.

In both stages of analysis, the incremental costs and QALYs for the intervention were calculated given the presence of a single or all of the capacity constraints. These incremental costs and QALYs were used to generate the incremental cost-effectiveness ratio (ICER) using the formula below:

\[
ICER = \frac{(Cost_2 - Cost_1)}{(QALY_2 - QALY_1)}
\]

(Equation 6.1)

The ICER estimated for the model in the presence of each capacity constraint and then all of them combined was compared against the NICE end-of-life threshold of £50,000 per QALY to determine whether the intervention was cost-effective.

As it is possible that capacity constraints will reduce the number of patients who receive the intervention, an additional output which incorporates the total number of patients is required. In this study, the static value of implementation method, which allows for varying marginal costs and benefits (see also Chapter 3) is used to quantify the impact of the capacity constraints. This method uses the total net benefit as the primary output of the analysis. The impact of a capacity constraint can be quantified by subtracting the total net benefit provided by introducing the intervention when the constraint, or all of the constraints, are in place from the total net benefit provided when no constraints are in place. This value is equivalent to the value of perfect implementation and is calculated using equation 6.2 below:

\[
\text{Value of Perfect Implementation} = n(NMB_{\alpha=1,\beta=1 or 0,\gamma=1} - p_{\alpha,\beta,\gamma} \cdot NMB_{\alpha,\beta,\gamma})
\]

(Equation 6.2)
Where \( n \) is the total population of patients who can be tested, \( \text{NMB}_{\alpha=1,\beta=1,\gamma=0} \) is the net monetary benefit per patient in the absence of the constraints, \( p_{\alpha,\beta,\gamma} \) is the proportion of patients receiving the intervention in the presence of the capacity constraints, and \( \text{NMB}_{\alpha,\beta,\gamma} \) is the net monetary benefit of the intervention in the presence of the constraints.

The variable \( \alpha \) is used to represent the lack of commissioning awareness constraint with a value of 1 meaning all patients are treated in hospitals who are aware of testing and a value of 0 meaning that no patients are treated in hospitals who are aware of testing.

The variable \( \beta \) represents the localisation or centralisation constraint. A value of 1 represents a situation in which all testing is offered through localised testing while a value of 0 represents a situation in which all testing is provided by centralised laboratories. As it is not known whether localisation of centralisation is the capacity constraint, equation 6.2 contains two potential values for \( \beta \) in the calculation of the NMB in the absence of capacity constraints. As such the value of \( \beta \) (1 or 0) which maximises this value will need to be found, with the other value representing the capacity constraint.

The variable \( \gamma \) represents the staffing level of pathology laboratories. A value of 1 represents a situation in which pathology laboratories are fully staffed. A value of 0 represents the level of pathology laboratory staffing in 2014.

Table 6.2 describes the value taken by each variable when determining the impact of each constraint individually or combined.
Table 6.2: Value for each capacity constraint in each scenario

<table>
<thead>
<tr>
<th>Capacity Constraint</th>
<th>Parameter Values</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>α</td>
</tr>
<tr>
<td>Lack of commissioning awareness</td>
<td>0.23</td>
</tr>
<tr>
<td>Localisation or centralisation</td>
<td>1</td>
</tr>
<tr>
<td>Insufficient pathology staffing</td>
<td>1</td>
</tr>
<tr>
<td>All constraints combined</td>
<td>0.23</td>
</tr>
</tbody>
</table>

α Proportion of patients at trusts who are aware of ALK testing commissioning arrangements
β Degree of localisation of ALK testing where 1 means all testing is localised and 0 means all testing is centralised
γ Ability of laboratories to conduct testing where 1 means there are no restrictions on laboratories ability to conduct testing and 0 means that there are restrictions

6.3.5.2 Sensitivity Analysis

Parameter uncertainty in the estimated size and effect on costs and QALYs of the capacity constraints was incorporated using probabilistic sensitivity analysis (276,277). Distributions were applied to the parameters and the analysis run 1000 times with new values drawn from the distributions in each iteration. This created a sample of 1000 estimates of the incremental cost and benefit, ICER, NMB, and NMB lost due to the presence of all of the capacity constraints.

The sample of incremental costs and QALYs estimated from the capacity constraint model PSA were plotted on an incremental cost-effectiveness plane. On the same figure, the distribution of incremental costs and QALYs resulting from the base line analysis (see section 5.3.2) was also plotted. The proportion of iterations with positive NMB was calculated to estimate the probability that the intervention would be cost-effective in the presence and absence of capacity constraints. The distributions of total NMB in the presence and absence of capacity constraints were contrasted
using kernel density plots. Kernel density estimation is a method for estimating the distribution of a
variable and kernel density plots present a visualisation of the estimated distribution (278).

6.3.5.3 Model Validation

To check the validity of the restructured model, all constraints were set to zero. In this situation the
model was expected to collapse into the baseline model created in chapter 5 and produce identical
results.

6.4 Results

This section presents the results of the base case analysis of: the impact of each capacity constraint
individual and combined; the sensitivity analysis; and the model validation

6.4.1 Base Case Analysis

Table 6.3 summarises the ICERs, NMB (current value of implementation), and NMB lost (value of
perfect implementation) due to the presence of each constraint individually and together. In the
first stage of the analysis, the ICER and impact on NMB of each capacity constraint individually
was evaluated. The lack of commissioning awareness constraint did not impact the ICER of the
intervention but resulted in a loss of NMB of £4,907,893. The localisation constraint resulted in an
intervention with a slightly higher ICER of £39,211 but a loss of NMB of only £7,773. Finally, the
pathology laboratory staffing constraint raised the ICER of the intervention to £40,322 and reduced
the NMB by £808,746.

To determine whether localisation or centralisation of testing was the capacity constraint, the
results produced when testing was fully localised or centralised were compared when no other
constraints were present. Full localisation of IHC testing yielded a marginally higher ICER
(£39,317 per QALY) and reduced net benefit (£6,303,222 per year) due to the small additional cost
of IHC testing in local laboratories. A comparison of the cost-effectiveness of the intervention
under full localisation or centralisation in the presence of the other constraints also suggested that
localised testing was the capacity constraint and that centralised testing yielded greater benefits for
society.
When combined, the three constraints resulted in an ICER of £41,413 and a loss of NMB of £5,289,414. This represents a loss of 83% of the potential benefit of the intervention.

### Table 6.3 The impact of the inclusion of capacity constraints on ICERs and NMB

<table>
<thead>
<tr>
<th>Capacity Constraints</th>
<th>Value of Constraint</th>
<th>ICER(^a)</th>
<th>Annual NMB(^b)</th>
<th>Annual Societal QALY gain</th>
<th>Annual NMB Shortfall Due to Constraint(^c)</th>
<th>Annual QALY Shortfall Due to Constraint(^d)</th>
</tr>
</thead>
<tbody>
<tr>
<td>None (Baseline)</td>
<td>100% of trusts aware of commissioning arrangements 0% of testing is localised/100% is centralised No impact of limited capacity on costs or outcomes</td>
<td>£39,198</td>
<td>£6,373,887</td>
<td>124.48</td>
<td>£0 (0%)</td>
<td>0 (0%)</td>
</tr>
<tr>
<td>Lack of awareness of ALK commissioning</td>
<td>23% of trusts aware how to commission tests</td>
<td>£39,198</td>
<td>£1,465,994</td>
<td>29.32</td>
<td>£4,907,893 (77%)</td>
<td>98.16 (77%)</td>
</tr>
<tr>
<td>Level of Localisation of Test</td>
<td>11% of testing done in house</td>
<td>£39,211</td>
<td>£6,366,114</td>
<td>127.32</td>
<td>£7,773 (0.1%)</td>
<td>0.16 (0.1%)</td>
</tr>
<tr>
<td>--------------------------------</td>
<td>--------------------------------</td>
<td>---------</td>
<td>------------</td>
<td>--------</td>
<td>----------------</td>
<td>-------------</td>
</tr>
<tr>
<td>Level of Pathology Lab Capacity</td>
<td>Impact of capacity on costs and outcomes of testing</td>
<td>£40,322</td>
<td>£5,565,141</td>
<td>111.30</td>
<td>£808,746 (13%)</td>
<td>16.17 (13%)</td>
</tr>
<tr>
<td>Expected Level of all Constraints in 2014</td>
<td>23% of trusts aware how to commission tests</td>
<td>£41,413</td>
<td>£1,084,473</td>
<td>21.69</td>
<td>£5,289,414^e</td>
<td>105.79^1 (83%)</td>
</tr>
<tr>
<td></td>
<td>27% of tests done in house</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>There is an impact of limited path capacity on costs and outcomes</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* ICER for ALK testing to guide crizotinib treatment with the health system capacity constraint in place compared to no testing and universal docetaxel

^ Net monetary benefit

^ NMB shortfall = NMB without constraints – NMB with one or all constraints present

^ QALY shortfall = QALYs without constraints – QALYs with one or all constraints present

^ The total NMB and QALY loss from the presence of all constraints is not a sum of that from individual constraints due to the interaction between localisation and pathology laboratory capacity
6.4.1 Sensitivity Analysis

The results of the probabilistic sensitivity analysis are shown in figure 6.2. This incremental cost-effectiveness plane features the plotted incremental costs and benefits calculated from 1000 Monte Carlo simulations of the cost-effectiveness of ALK testing and crizotinib in the presence and absence of all three selected capacity constraints. The expected incremental costs and QALYs in the decision-analytic model including capacity constraints are lower as only a relatively small proportion of patients receive testing, and subsequent crizotinib, due to the lack of commissioning awareness in pathology laboratories. The effect of the localisation and pathology capacity constraints is to reduce the probability of cost effectiveness from 86% to 78%, showing there is an increase in the observed uncertainty in the model (279).

Figure 6.2 Incremental Cost-Effectiveness Plane for the Results of the Model in the Presence and Absence of Capacity Constraints

Footnote: incremental QALYs plotted against incremental costs of ALK testing to guide crizotinib compared to no testing when capacity constraints are present (orange points) or not present (black points).

Figure 6.3 shows the distributions of the estimated total net monetary benefit of ALK testing and crizotinib. While the decision-analytic model including capacity constraints yields a lower
expected NMB and has more iterations with NMB below zero, there is less uncertainty in the estimates. In the absence of constraints there is a higher NMB and fewer iterations with an NMB below zero but a much larger distribution of potential values. In some instances, it is possible that the NMB for the intervention in the absence of constraints is significantly lower than the NMB in the presence of constraints.

Figure 6.3 Kernel Density Plot of Estimated Total NMB in the Presence and Absence of Capacity Constraints

6.4.2 Model Validation

To determine the face validity of the model, all of the capacity constraints were turned off in the model to ensure that the same results were generated as in the baseline model. Setting the commissioning and pathology constraints to zero and offering fully centralised IHC testing yielded the same ICER (£39,198 per QALY) and NMB (£6,373,887 per year) figures as the baseline model
(see Chapter 5) indicating that no errors in the underlying evaluation had been created by restructuring the decision model.

6.5 Discussion

This study investigated the impact of incorporating capacity constraints in a decision-analytic model which sought to estimate the cost-effectiveness of ALK testing and crizotinib as an example of precision medicine in NSCLC. The inclusion of three capacity constraints, which were selected from the 17 barriers identified in in qualitative interviews with stakeholders (Chapter 4), had the effect of reducing the cost-effectiveness of ALK testing and crizotinib and limiting the amount of net societal benefit its introduction produced for the health system.

The results of the analysis of the impact of each individual capacity constraint suggested that constraints could be grouped into two main categories based on their impact on the cost-effectiveness or net societal benefit of the intervention. In this study these categories of constraints will be referred to as “access limiting” or “quality limiting”. In this study the commissioning awareness constraint was an example of an “access limiting constraint”. This constraint limited the number of patients who could receive testing and while this significantly limited the net societal benefit of the intervention, it did not reduce its cost-effectiveness. This was because in this case study the incremental costs and QALYs of ALK testing and crizotinib did not depend on the number of patients receiving the test or treatment, meaning that marginal costs and benefits were constant. However, if the marginal costs and benefits of ALK testing and crizotinib did depend on the number of patients receiving it, as in the case study presented in chapter 3, then this type of “access limiting” constraint would also impact the cost-effectiveness of the test-treat intervention.

An “access limiting” constraint can therefore be defined as a constraint which reduces the number of patients who receive a potentially cost-effective intervention resulting in a definite change in NMB and a potential change in the ICER depending on whether incremental marginal costs and QALYs are fixed or variable. The presence of access limiting capacity constraints which do not impact the ICER highlights the importance of going beyond the use of ICERs and averaged net benefit calculations as the sole outcome of economic evaluation when assessing the impact of capacity constraints. Total population level NMB calculations are needed to capture the impact of these constraints.
The other type of capacity constraints are “quality limiting” capacity constraints. The localisation versus centralisation and pathology laboratory capacity constraints in this case study were examples of this type of constraint. The impact of these constraints was to either raise the estimated incremental cost or lower the estimated incremental benefit per patient receiving the intervention. Contrary to “access limiting” constraints, the presence of these constraints directly impacts the cost-effectiveness of ALK testing and crizotinib as measured by the ICER. Furthermore, as the incremental costs and QALYs are key to the determination of net benefit and net benefit is a key component of societal net benefit, “quality limiting” constraints also reduce the NMB of the intervention. This can be seen in the estimated individual impacts of the localisation versus centralisation and pathology laboratory capacity constraints in this case study. In this case study, the impact of the “quality limiting” capacity constraints on the estimated ICERs were relatively small. However, this still translated to a loss of QALYs (16) in the case of limited pathology laboratory capacity.

When combined, the three capacity constraints caused a loss of NMB of £5,289,414 which equates to approximately 106 QALYs at the end-of-life threshold. Implementing ALK testing and crizotinib in 2014 without addressing any of the constraints would have only yielded 17% of the potential net societal benefit that the test-treat intervention could deliver. While in this case ALK testing and crizotinib was still cost-effective and therefore yielded positive net societal benefits even in the presence of constraints, it is conceivable that quality limiting capacity constraints could push the ICER for the test-treat intervention above the threshold, resulting in a net societal loss. In this case, reducing “access limiting” constraints without addressing the “quality limiting” constraints could increase this net societal loss by increasing the provision of a cost-ineffective intervention. This issue is considered further in chapter 7.

The interaction between the types of constraints is apparent in the kernel density plots for the estimated samples of total societal net benefit created in the probabilistic sensitivity analysis. In the decision-analytic model including capacity constraints a greater proportion of the distribution of results lies below zero, indicating a lower probability of cost-effectiveness of ALK testing and crizotinib. However, the impact of this is mitigated by the small proportion of patients who receive ALK testing due to a lack of commissioning awareness. This means that when there is a net societal loss, this is relatively small as few patients receive ALK testing and crizotinib. When there are no constraints there is a greater chance that ALK testing and crizotinib is cost-effective but there is also the potential for a significant net societal loss.
Nine published studies incorporating capacity constraints in economic evaluations were identified in a systematic review (see Chapter 2). There were some similarities with regards to the impact of capacity constraints between the findings of this study and the findings of the nine studies identified in the systematic review. The study by Retèl et al., (2012) investigating the impact of barriers and facilitators to introducing a Mammaprint test for breast cancer provides a good example of how “access limiting” constraints interact with varying marginal costs and benefits (136). In this example (see also Chapter 3), Mammaprint has high initial marginal incremental costs and low marginal incremental benefits. The access limiting constraints of uptake by clinicians, non-compliance with results, and the failure rate of the test acted to limit the amount of patients who received Mammaprint meaning that the ICER for the technology remained high. Over time the impact of these constraints reduced, increasing the ICER and eventually making the Mammaprint test cost-effective.

Examples of “quality limiting” constraints were present in a 3 studies (139,153,155) identified in the systematic review (see Chapter 2). In a study of the impact of adherence to tamoxifen by patients with breast cancer, McCowan et al., found that patients with an adherence of less than 80% were expected to lose 1.12 QALYs and experience an additional £5,970 in medical costs compared to patients with over 80% adherence (153). In two studies, the inclusion of quality limiting constraints had a significant impact on the decision as to which treatment to offer (139,155). In a study on NSCLC, when the turnaround time for a multiplexed biomarker test was increased 1.5 fold, the optimal strategy changed from test and then treat to beginning standard chemotherapy before test results were returned (139). In an economic evaluation of the cost-effectiveness of letrozole versus tamoxifen for women with breast cancer, letrozole was a dominant option when drug wastage of tamoxifen was 15% but had a positive incremental cost when drug wastage was 0% (155).

While the test-treat intervention in this case study still created positive net benefits for society, it is possible that for other interventions the inclusion of capacity constraints will result in the intervention being deemed not cost-effective at the time of evaluation. This may have significant implications for the evaluation of examples of precision medicines by technology appraisal agencies such as NICE. Currently, such constraints are not included in economic evaluations and so the results of such studies should arguably be taken as long-run estimates of cost-effectiveness (109,111,113).
Failing to account for the need for capacity investments may also have significant implications for strategic drug pricing. Strategic drug pricing is the practice of drug manufacturers setting a price for their medicine such that it is narrowly deemed cost-effective. If appraisal decisions are made based on long-run estimates of the cost-effectiveness of an intervention then the cost of any investments required to make the intervention cost-effective or improve access to the intervention will fall on the health system rather than being born by the drug manufacturer in terms of a reduced price of the drug. Including the cost of any required capacity investments in the HTA will reduce the maximum price the drug manufacturer can charge for their medicine while keeping it cost-effective.

Limitations

There were some limitations to this study that mainly centred around the lack of available data. A significant issue was that there was limited information available about the number of ALK tests offered in 2014, the impact of delaying treatment initiation due to a long turnaround time for tests, and the impact of limited pathology laboratory capacity. As such, a number of assumptions had to be made about these values. In addition, it was difficult to determine the mechanism by which the capacity constraints impacted the costs, QALYs, and level of provision of ALK testing and crizotinib. For example, it was hypothesised that longer turnaround times for ALK testing would result in poorer treatment outcomes for patients but due to confounding with severity of disease, applied studies found the opposite to be true. To determine the impact of this constraint the researcher contacted two consultant oncologists who believed that turnaround times of two-weeks would not lead to worse treatment outcomes for patients but would lead to an additional visit with an oncologist and an anxiety related disutility. In future applied studies, more robust methods should be used to determine the mechanisms by which constraints impact cost-effectiveness and the resulting parameter values. This could involve the use of focus groups, the Delphi method, or expert elicitation (280,281).

This study selected three capacity constraints based on the results of a qualitative study (see Chapter 4). This qualitative study sought to identify the range of barriers to using examples of precision medicine in non-small cell lung cancer. From these constraints, three were selected for the case study model on ALK testing and targeted treatment because they were commonly mentioned and had a clear mechanism for how they may impact on cost-effectiveness and NMB.
However, this does not necessarily mean that these constraints were the ones which had the greatest impact on these outcomes in practice. This may be a particular problem for future prospective studies which seek to incorporate capacity constraints. If many participants anticipate a particular barrier then they may be prepared to take measures to overcome it when faced with it in practice. More damaging may be the unforeseen constraints which may take longer to overcome. It is key that the most important capacity constraints are selected to be including in a prospective decision analytic model based economic evaluation. In order to achieve this, more appropriate methods such as the Delphi method should be used with stakeholders in the implementation process to understand the degree of consensus within a sample of experts about which constraints may have the largest impact and should be included.

The assumed value used for the cost-effectiveness threshold may also represent a limitation in terms of the impact on the cost-effectiveness of ALK testing and crizotinib. The results of the quantification of the impact of capacity constraints should be put in context. The end-of-life threshold used in this case study is argued to represent a demand side threshold representing the value placed on the gain of a QALY for patients with terminal illness by the public (282). This contrasts with the alternative theoretical view of the threshold in that it represents the marginal cost-effectiveness of the health system. The use of the end-of-life threshold means that the values for net monetary benefit and net health benefit actually represent net preference weighted monetary benefit and net preference weighted health benefit.

Threshold values provided for the marginal cost-effectiveness of the system tend to be lower, either in the form of the main NICE threshold of £20 to £30,000 per QALY or the empirical estimate for the threshold of approximately £13,000 per QALY provided by researchers at the University of York (283). Given that the ICER estimated for the intervention in the case study was £39,198 per QALY, the intervention is cost-effective at the demand-based threshold and not cost-effective at either supply-based threshold. Implementing the intervention therefore increases the number of preference weighted QALYs provided by the system but decreases the absolute number of QALYs provided. The estimates of the NMB or QALYs lost due to capacity constraints should also be interpreted as preference weighted losses.

6.6 Conclusion
This study has shown that the inclusion of health system capacity constraints reduced the estimated cost-effectiveness and societal net benefit of ALK testing and crizotinib. This means that economic evaluations which omit capacity constraints may give misleading representations of the cost-effectiveness of examples of precision medicine. In addition, including measures of total net benefit in economic evaluations which account for capacity constraints may help to highlight situations in which the intervention will not be available to all patients at the time of approval whereas the ICER does not contain this information.

6.7 Summary of Chapter 6

Chapter 6 described the selection of three key constraints that were faced in the introduction of ALK testing to guide treatment with crizotinib and described how a traditional decision-analytic model based economic evaluation (see Chapter 5) could be adapted to incorporate these constraints. The static value of implementation method developed in chapter 3 was used to determine the total net benefit lost from the presence of each constraint separately and when they were combined.

Chapter 7 will explore the value of investing in improving health system capacity to reduce the impact of capacity constraints and realise increased benefits for patients and the health system. A hypothetical capacity investment to address each of the constraints identified in this chapter will be described. The dynamic value of implementation method, allowing for varying marginal costs and benefits, developed in chapter 3 will be used to determine whether each capacity investment represents a cost-effective use of resources.
Chapter 7
Quantifying the Value of Investments to Reduce the Impact of Capacity Constraints in Precision Medicine

7.1 Chapter 7 Summary

In chapter 6, a method was presented to estimate the one-time (static) impact of capacity constraints to an example of precision medicine at the time of evaluation by a national decision-making body (NICE). It was shown that consideration of capacity constraints may result in a shortfall of the net societal monetary benefit, or health gain, when compared with the perfect uptake and use of an example of precision medicine (ALK testing and crizotinib). It was shown that the presence of two of the capacity constraints could impact on the marginal cost-effectiveness of the intervention.

Reducing the impact of capacity constraints, by investing in specific strategies, could potentially improve the cost-effectiveness and net monetary benefit produced by ALK testing and crizotinib. Such investment has a cost and funding for the strategies to reduce the impact of capacity constraints must be taken from elsewhere in the health system. It is therefore important to investigate whether the benefits of such investments outweigh the benefits foregone by other patients using appropriate methods. Chapter 7 reports a study that uses dynamic value of implementation methods to assess the value of investing in overcoming capacity constraints to providing a case study example of precision medicine: ALK testing to guide crizotinib.

7.2 Introduction

A study that used static value of implementation methods (see Chapter 6) showed the potential impact of capacity constraints on the introduction of an example of precision medicine. The results of this analysis suggested that the three identified capacity constraints may have had a significant impact on the cost-effectiveness and net monetary benefit of the ALK testing and crizotinib. The inclusion of all three capacity constraints in the decision analytic model reduced the total net monetary benefit of the intervention by 83%.
The reduced NMB of *ALK* testing and crizotinib suggested that there may value in undertaking strategies to reduce the impact of the constraints and increase the benefit for patients with NSCLC. These implementation strategies (capacity-investments), will potentially yield benefit in improving the net monetary benefit of the test-treat intervention but will themselves have a cost. As the resources needed to fund these capacity-investments will be taken from other interventions in the health system, it is important to evaluate whether the gain in outcomes for patients with NSCLC will be greater than the outcomes lost by those losing funding.

Chapter 3 developed dynamic value of implementation methods to determine the value of such capacity-investments when marginal costs and benefits vary. As was shown in chapter 6, the localisation and pathology laboratory staffing constraints both impacted the marginal costs and QALYs of *ALK* testing and crizotinib and therefore the methods outlined in chapter 3 will be required to evaluate investments in reducing these constraints.

To date value of implementation methods have not been used in the field of precision medicine (see Chapter 2). Furthermore, value of implementation methods have not been used to address the value of investing in improving health system capacity using capacity-investments.

### 7.2.1 Aims and Objectives

The aim of this study was to investigate the value of a defined number of capacity-investments to improve the ability of a constrained health system to provide *ALK* testing and targeted treatment with crizotinib to patients with NSCLC.

In order to achieve this aim, this study had six objectives to:

- Convert a static decision-analytic model-based cost-effectiveness analysis of *ALK* testing and crizotinib (see Chapter 6) into a dynamic decision-analytic model-based cost-effectiveness analysis;
- estimate the growth in the population of patients who may benefit from *ALK* testing in future time periods;
- estimate the natural diffusion of *ALK* testing in the absence of any capacity-investments;
- identify and parameterise an exemplar capacity-investment for each capacity constraint;
- evaluate the value in terms of improving the net societal benefit of *ALK* testing and crizotinib of each capacity-investment in isolation and all capacity-investments together;
• explore whether the value of the capacity-investments depends on whether ALK testing and crizotinib is cost-effective at the time of its entrance into the health system.

7.3 Methods

This study used a dynamic value of implementation approach, allowing for varying marginal costs and benefits, to adapt an existing decision-analytic model-based cost-effectiveness analysis (see Chapter 5) of ALK mutation testing to guide treatment with crizotinib compared with universal docetaxel to quantify the impact of defined capacity-investments to reduce the impact of three selected capacity constraints (see Chapter 6). Unless otherwise specified in this chapter, the key characteristics of the model such as the study perspective (NHS), price year (2019) and discount rate (3.5%) are the same as those reported in the baseline model in chapter 5 and static model in chapter 6. The time horizon for this dynamic analysis was assumed to be six years.

7.3.1 Dynamic Impact of Capacity Constraints: Key Steps

To implement this dynamic approach five steps are required:

1. convert the existing decision-analytic model (see Chapter 6) into a dynamic model;
2. define and parameterise the implementation strategies to improve the capacity of the health system to deliver the intervention;
3. identify the expected patient population size for each year;
4. identify the expected pattern of natural diffusion of the intervention in the absence of strategies to improve its implementation;
5. run the model for each time period with parameters updated in each year to reflect changes in implementation, costs of implementation strategies, natural diffusion, and patient population size.

7.3.2 Converting the Static Decision Analytic Model into a Dynamic Model

Dynamic value of implementation analysis is used to evaluate the value of implementation strategies that improve the use of a cost-effective intervention over time. While traditional economic evaluations evaluate the cost-effectiveness of an intervention for a single cohort of
patients, dynamic models include multiple cohorts of patients over many years. In future years, key model parameters may vary, altering the cost-effectiveness and value of the intervention.

To convert the static decision analytic model into a dynamic model, a loop was created in R. In this instance, rather than running the static analysis multiple times with new parameter values drawn from distributions to represent uncertainty, as would be done in a probabilistic sensitivity analysis, new values were used to represent the change in the parameter values over time. For example, as is outlined in section 7.2.5, the patient population may grow over time. In the first year to be evaluated an initial value for the patient population size is used in the decision-analytic model and model-outputs stored. In the second iteration of the run of the decision-analytic model, the population size variable is replaced by a new value representing the increased patient population in the next year and the model-outputs stored. This process continues for each year, using new values for each parameter which varies over time.

The assumed time horizon for this dynamic model of six-years was chosen to represent the time from the initial appraisal of crizotinib by NICE to the present day (2019). It was assumed that the decision about whether to invest or not is taken at the beginning of a year and the model runs in cycles of one year.

7.3.3 Specifying Investment Strategies to Reduce Capacity Constraints

This study defined three capacity-investment strategies to target the three capacity constraints that were identified in the study reported in chapter 4. The targeted capacity constraints were: a lack of awareness of how ALK testing is commissioned, localised ALK testing, and insufficient pathology laboratory staffing. The capacity-investment strategies evaluated in this chapter are based on the qualitative interviews in chapter 4 and evidence from published literature. These are capacity-investments are now described.

Capacity Constraint 1: Awareness of commissioning arrangements for ALK mutation testing

The lack of awareness of commissioning arrangements for ALK mutation testing meant that a large proportion of the patient population may not receive testing and so miss out on potentially life extending and improving targeted treatment (crizotinib). While awareness of commissioning arrangements may naturally increase over time, as modelled by a diffusion curve (see figure 7.2), the use of implementation strategies could increase the speed of diffusion and modify the shape of the curve.
The capacity-investment for this constraint was identified from a published value of implementation study which sought to value strategies to improve the adherence of clinicians and cancer survivors to physical exercise guidelines (165). Five capacity-investment strategies targeting clinicians were selected from this list: continuing education meetings and workshops; educational outreach visits; use of printed educational materials; enlisting local opinion leaders; instituting audit and feedback mechanisms and reminder systems. The potential effects of the implementation strategies were identified by Mewes et al (2017) from published systematic reviews (165). The associated resource use was calculated by Mewes et al and included “how much time was required by each type of staff for training and for executing the strategy, and which material, services, and/or travel and meeting costs were needed”. Costs were applied using the Dutch manual for cost research. Four of these strategies, that may be applicable to learning about guidelines as opposed to adhering to them, were selected as relevant for a capacity-investment strategy applicable to a one-off event of learning about the commissioning arrangements for ALK testing (see table 7.2)

Mewes et al (2017) identified the direct medical costs of the capacity-investment and these were separated into development costs and a cost of targeting the strategy to a population of clinicians (per staff member). These costs were converted from euros to pounds using the European Central Bank Statistical Data Warehouse (284) and deflated to 2014 values using an averaged Hospital and Community Health Services index value between 2014 and 2017 (0.83% per annum) (285). The costs were then reduced to reflect that 23% of patients were already receiving ALK testing so the strategy would not be needed in 23% of hospitals. It was assumed that the development costs (see table 7.2) would therefore be 77% of that outlined in Mewes et al., (2017). It was assumed that each capacity-investment would only be required to reach one health professional in each hospital that did not currently offer testing (136) and that patients were distributed evenly across hospitals. While the source study measured a static impact of the capacity-investment, in this study it was assumed that the effect was sustained across six-years (see table 7.4). From these potential implementation strategies, educational outreach visits were chosen as the example capacity-investment to use in this analysis because the strategy exhibits the lowest cost per percentage increase in provision of ALK testing.
### Table 7.1: Cost and effect of the capacity-investment to Improve Awareness of ALK test commissioning

<table>
<thead>
<tr>
<th>Strategy</th>
<th>Description</th>
<th>Cost (£; 2017)</th>
<th>Proportional increase in ALK testing provision</th>
</tr>
</thead>
<tbody>
<tr>
<td>Enlisting local opinion leaders</td>
<td>Identify relevant opinion leaders who have a “unique and influential position in their system’s communication structure” (286). Communicate commissioning arrangements to these individuals who have a large informal reach and are seen as trustworthy. Opinion leaders hold tutorials with hospitals in their region about reimbursement for ALK testing</td>
<td>Development Costs: £533,818</td>
<td>9%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Cost per hospital: £1,064</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Total Cost: £678,522</td>
<td></td>
</tr>
<tr>
<td>Continuing Education Meetings/Workshops</td>
<td>NHS staff are often required to attend meetings and undertake training to keep up to date with current practice. This strategy would see the commissioning</td>
<td>Development cost: £26,001</td>
<td>6%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Cost per hospital: £888</td>
<td></td>
</tr>
<tr>
<td>Services</td>
<td>Description</td>
<td>Total Cost</td>
<td>Percentage</td>
</tr>
<tr>
<td>--------------------------------------</td>
<td>--------------------------------------------------------------------------------------------------------</td>
<td>-----------------------------</td>
<td>------------</td>
</tr>
<tr>
<td>Educational Outreach Visits</td>
<td>Staff are visited in their own premises by a person with knowledge of the commissioning arrangements for ALK testing</td>
<td>Development costs: £13,048</td>
<td>5%</td>
</tr>
<tr>
<td></td>
<td>Cost per hospital: £787</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Total cost: £120,080</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Printed Educational Materials</td>
<td>Information about the commissioning arrangements is sent to relevant staff. In this example three information emails are developed by a researcher and sent three times using mailing lists</td>
<td>Development costs: £600,992</td>
<td>2%</td>
</tr>
<tr>
<td></td>
<td>Cost per hospital: £15.47</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Total Cost: £603,096</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Capacity Constraint 2: Localisation of ALK testing**

The second relevant capacity constraint was the degree of localisation or centralisation of testing. In the qualitative interviews in chapter 4 there had been disagreement between participants about whether testing should be done in local laboratories in each hospital to help produce faster turnaround times for tests or larger centralised laboratories who had more expertise in testing but potentially slower turnaround times. It was found in chapter 6 that localised testing provided worse outcomes due to the lower quality of testing. This finding means that a policy of centralising ALK IHC testing in pathology labs may be valuable. As such, the value of a capacity-investment to influence centralisation and quality assurance of ALK IHC testing was considered in this study.
In chapter 6, it was estimated that 27% of hospitals offered ALK testing in-house in 2014. Given that there were 176 hospitals this meant that ALK testing was conducted locally in approximately 48 hospitals. Under a policy of centralisation (see also section 6.2.4.2) there will be an optimum number of pathology laboratories offering testing. Given the recent move to consolidate pathology laboratories into 29 networks in the UK, it was assumed that one laboratory in each network hub would offer ALK testing in an ideal scenario (214). This assumption meant that this capacity-investment would require 19 laboratories to transfer their testing to a central laboratory.

It has been suggested that the move from individual pathology laboratories to networks will save the NHS £200 million between 2016 and 2021. The potential savings from centralising ALK testing would be very small compared with this figure but would still represent a cost saving. However, this would be captured in the model by the lower cost of IHC testing for central laboratories and so including the cost saving of the capacity-investment would be double counting. Instead it is assumed that the 19 hospitals offering localised testing would need some initial convincing that centralising testing would be beneficial to their patients. In this hypothetical example it is assumed that to centralise testing fully, each of the 29 laboratories would have to offer a validated ALK test and that each of the localised laboratories would receive an educational outreach visit as outlined in the “awareness of commissioning arrangements for ALK mutation testing” constraint. Participants in the interview study reported in chapter 4 suggested that validation of ALK testing cost £5,000 per laboratory, suggesting a total of £145,000 to validate testing in all of the 29 laboratories. The educational outreach visits would have a development cost of £13,048 and a cost of visiting the 19 localised testing hospitals of £14,953. The total cost of this hypothetical implementation strategy was therefore £173,001. It was assumed that this would result in full centralisation of ALK testing a year after approval of crizotinib.

**Capacity Constraint 3: Under-staffing of pathology laboratories**

The observed higher costs and lower QALYs associated with both centralised and localised testing (see Chapter 6) compared with the results of the baseline decision-analytic model (see Chapter 5) were a product of under-staffed pathology laboratories. Centralised testing may yield improved outcomes compared with localised testing but centralisation will not result in perfect implementation of testing unless the staffing of the centralised laboratories is also improved. This link highlights the direct interaction between these two capacity constraints.
It was assumed that pathology laboratory capacity can be improved by hiring additional pathologists as this was highlighted as an issue by pathologists in the qualitative interview study. The report on pathology laboratory capacity undertaken by cancer research UK suggested that of the hospitals responding to a given survey question 20 out of 36 (56%) had an unfilled vacancy for at least one full time equivalent pathology consultant and 32 out of 42 (76%) respondents had at least one vacancy for a biomedical scientist (213).

The proportion of hospitals with vacancies for these professions was applied to the number of hospitals offering *ALK* testing in the dynamic model. In the base case this was set at 48 vacancies but falls to 29 vacancies if testing is centralised meaning that two cost estimates must be determined. It was assumed that filling all of the available vacancies will remove the negative effects of centralisation or localisation. It was also assumed that filling a role will reduce these negative effects proportionate to the number of total vacancies. For example, if the biomedical scientist vacancies can be filled after three-years then this will reduce the negative effects of the centralisation of localisation strategies by 58% as this is the proportion of total vacancies which require biomedical scientists. Later filling the consultant positions after five-years have passed will remove the remaining 42% of negative effects.

Table 7.2 reports the unit costs for hiring pathology consultants and biomedical scientists. These unit costs were used to calculate the total costs for the base case (partly localised testing) and centralised testing assumption (see table 7.3). In calculating the cost of additional pathologists and biomedical scientists it was assumed that the costs of recruitment and training where included in the administrative support and training cost components reported in the PSSRU Unit Costs (272) for these roles. However, this does raise an important issue in that it may take a substantial amount of time to train additional staff if there is not a sufficient supply in the current workforce. A biomedical science degree takes three-years to complete and from completion of foundation doctor training it takes five-years to train to become a consultant pathologist. As such while the policy of increasing pathology laboratory staffing will be considered to start in 2014, the costs and effects of increased biomedical scientist staffing will only be incorporated into the model from 2017 while those of increased consultant staffing will only be included in 2019 which is the last year in the model.
### Table 7.2: Unit staff costs

<table>
<thead>
<tr>
<th>Staff Type</th>
<th>Unit Cost (per year; 2014)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Consultant(^1)</td>
<td>£254,819</td>
</tr>
<tr>
<td>Biomedical Scientist(^2)</td>
<td>£83,599</td>
</tr>
</tbody>
</table>

\(^1\) Based on cost of a medical consultant from PSSRU (2014)

\(^2\) Based on costs of an average grade 6 hospital post from PSSRU (2016) (deflated to 2014 figures)

### Table 7.3: The Cost of Filling Vacancies in Pathology Laboratories with Different Levels of Centralisation

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Base Case (partly localised testing)</th>
<th>Centralised Testing</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of Laboratories</td>
<td>48</td>
<td>29</td>
</tr>
<tr>
<td>Number of Consultant Pathologists Required</td>
<td>27</td>
<td>16</td>
</tr>
<tr>
<td>Number of Biomedical Scientists Required</td>
<td>37</td>
<td>22</td>
</tr>
<tr>
<td>Total Annual Cost of Additional Consultant Pathologists</td>
<td>£6,880,113</td>
<td>£4,077,104</td>
</tr>
<tr>
<td>Total Annual Cost of Additional Biomedical Scientists</td>
<td>£3,093,145</td>
<td>£1,838,298</td>
</tr>
<tr>
<td>Total Annual Cost of Filling Pathology Vacancies</td>
<td>£9,973,258</td>
<td>£5,915,402</td>
</tr>
<tr>
<td>Proportion of Time Allocated to ALK testing(^1)</td>
<td>0.28%</td>
<td>0.28%</td>
</tr>
<tr>
<td>Total Annual Cost of Recruiting Additional Pathologists and Biomedical Scientists to Conduct ALK Testing</td>
<td>Consultants: £19,264 per year</td>
<td>Consultants: £11,416 per year</td>
</tr>
<tr>
<td></td>
<td>Biomedical Scientists: £8,661 per year</td>
<td>Biomedical Scientists: £5,147 per year</td>
</tr>
<tr>
<td></td>
<td>Total: £27,925 per year</td>
<td>Total: £16,563 per year</td>
</tr>
</tbody>
</table>

\(^1\) Predicted number of ALK tests (16,627) as a proportion of total histopathology requests (33,454 per lab times 176 hospitals), Cancer Research UK (2016), National Lung Cancer Audit (2014)
7.3.4 Defining the Relevant Patient Cohort

As the effects of the capacity-investments will occur over a number of time periods, they are assumed to affect future cohorts of patients rather than just the initial cohort of patients that receive *ALK* testing and crizotinib at the time of approval. It is likely that the population of patients who could receive *ALK* testing and crizotinib will change in size in future years and as value of implementation analysis uses total net monetary benefit as an measure of output, changes in the patient population size will impact on the value of implementation.

In order to determine the number of patients eligible for *ALK* mutation testing, the total number of new lung cancer cases in each year was extracted from the National Lung Cancer Audit (264,287–289). As these audit data are only available up to 2017, a regression-analysis, using ordinary least squares, conducted in Microsoft Excel (106) was used to estimate the total number of cases for 2018 and 2019:

\[ y = 33,592 + 2155\text{ypa} \]  
(Equation 7.1)

Where \( y \) = the number of patients with lung cancer and \( \text{ypa} \) = the number of years after the potential approval of crizotinib in 2014. The predicted fit of this model was reasonable with a \( R^2 \) of 0.91.

The estimated number of new NSCLC cases was available in 2014 and this was used to create an estimate of the proportion of lung cancer cases which were NSCLC (68%). This value was then applied to the total case numbers for 2015 to 2019. The method was used to predict the proportion of cancers which were stage III or IV using figures presented by the manufacturer in the NICE TA296 submission (234). Finally, 15% of patients were removed from the eligible sample because this proportion would have been expected to receive positive *EGFR* results. Table 7.4 shows the estimated population size and the predicted number of patients who would receive crizotinib with full implementation of testing.
Table 7.4: Predicted Test Population Size in Future Years

<table>
<thead>
<tr>
<th>Year</th>
<th>New Lung Cancer Cases</th>
<th>NSCLC Cases</th>
<th>Stage III or IV NSCLC Cases</th>
<th>EGFR Stage III or IV NSCLC Cases (Test Population)</th>
<th>EGFR-, ALK+ Stage III or IV NSCLC (Crizotinib Population)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2014</td>
<td>33,027</td>
<td>22,461</td>
<td>19,561</td>
<td>16,627</td>
<td>565</td>
</tr>
<tr>
<td>2015</td>
<td>36,025</td>
<td>24,500</td>
<td>21,337</td>
<td>18,136</td>
<td>617</td>
</tr>
<tr>
<td>2016</td>
<td>39,041</td>
<td>26,551</td>
<td>23,123</td>
<td>19,654</td>
<td>668</td>
</tr>
<tr>
<td>2017</td>
<td>39,205</td>
<td>26,663</td>
<td>23,220</td>
<td>19,737</td>
<td>671</td>
</tr>
<tr>
<td>2018</td>
<td>41,360</td>
<td>28,128</td>
<td>24,496</td>
<td>20,822</td>
<td>708</td>
</tr>
<tr>
<td>2019</td>
<td>43,515</td>
<td>29,594</td>
<td>25,773</td>
<td>21,907</td>
<td>745</td>
</tr>
</tbody>
</table>

7.3.5 Diffusion

Diffusion is defined as “a passive phenomenon of adoption by individuals and organisations” (290). Precision medicine may naturally diffuse, with no adaptive strategies to modify uptake and use, into clinical practice. If the impact of diffusion is omitted from decision-analytic model-based cost-effectiveness analysis, then the value of any capacity-investments which aim to actively promote implementation of the intervention may be overestimated.

The course of diffusion of health technologies is generally shown using a diffusion curve. Such curves typically follow an “S-shape” with initially slow uptake by early adopters. The diffusion of the intervention then increases in pace before slowing again due to the presence of a small number of hesitant or resistant adopters. Figure 7.5 shows some examples of diffusion curves for a preterm birth screening technology which have been reproduced from a published paper (291). The specific shape of the curve may vary to a number of factors and previous research has identified difficulties in synthesising such curves due to heterogeneity which may be more qualitative than quantitative (292). In the example presented in Figure 7.5, different experts had different opinions about the speed at which the technology would diffuse into practice and this is represented by changes in the steepness of the diffusion curves.
There are very limited data (see Chapter 6, section 6.2.4.2) regarding the diffusion of ALK testing in UK clinical practice (265). Due to the limited availability of data, a pragmatic approach was taken and it was assumed that the rate of diffusion for ALK testing would be similar to that of EGFR testing. A 2015 report by Cancer Research UK contains a graph displaying the number of EGFR tests provided compared with the potential demand for these tests (99). This information was based on information requests sent to UK laboratories. As no individual level data were available, the software Digitizeit (250) was used to read published graphs (99) and provide readings of the level of provision of tests in specified years between June 2010 and June 2014. These levels of test use, alongside the estimated demand for tests, were used to populate a simulated data set of individual patients who could potentially have received the test over the specified time-frame. Each simulated patient was assigned to a year and a variable was created representing whether the patient had received testing or not with the proportion of patients receiving testing in each year.

Using this simulated dataset, a logistic regression was then conducted in the statistical software package R (224). The dependent variable was defined as receipt of testing (yes/no) and year post approval (ypa) of gefitinib (2010) as the explanatory variable. A constant term was included to represent the probability of a patient receiving an EGFR test in 2010.

\[
\text{Probability of EGFR testing being available} = \frac{e^{(-1.870+1.035ypa)}}{e^{(-1.870+1.035ypa)} + 1} \quad (\text{Equation 7.2})
\]

Equation 7.2 was used to predict the probability that a patient would have access to EGFR testing in a specified year between 2010 and 2015. The data taken from the report published by Cancer Research UK (99) did not include the proportion of patients who had access to EGFR testing.
immediately on approval of the first EGFR TKI. An estimate of 13% of patients was predicted using the logistic regression. While EGFR testing would not have been widely demanded before the approval of gefitinib, some hospitals may have been offering the treatment or may have been involved in trials and so some level of testing before the introduction of gefitinib is a reasonable assumption.

These uptake figures were adapted to predict the uptake figures for ALK testing. The baseline availability of testing was modified to 23% of patients. This assumption was made based on the analysis reported in chapter 6, section 6.3 that suggested awareness of commissioning arrangements would be a key constraint acting on the access to ALK testing at the time of evaluation of crizotinib. To do this, the constant term (see equation 7.2) to predict ALK testing provision was set to -1.208 such that in 2014 when crizotinib was being evaluated by NICE, 23% of patients were receiving ALK testing. Figure 7.3 shows the predicted availability of ALK and EGFR mutation testing based on the predictions of the logistics regression.

\[
\text{Probability of ALK testing being available} = \frac{e^{(-1.208+1.035y_{pa})}}{e^{(-1.208+1.035y_{pa})} + 1} \quad \text{(Equation 7.3)}
\]

Figure 7.2: Natural diffusion of EGFR and ALK testing
7.3.6 Analysis

To determine the economic value of the three proposed capacity-investment strategies, the dynamic incremental net monetary benefit of implementation (see Chapter 3, sections 3.2.4) was calculated using the decision-analytic model-based cost-effectiveness analysis of ALK testing to guide crizotinib compared with universal docetaxel (see Chapter 5). Equation 7.3 shows how the varying marginal incremental costs and benefits were calculated to generate a measure of incremental net monetary benefit:

\[
\text{INB of Implementation} = \sum_{t=1}^{6} \frac{n_t(\sigma_t NMB_{\beta,y} - \alpha_t NMB_{\beta=0.11,y=1}) - i_t}{(1 + r)^{t-1}} \quad \text{(Equation 7.4)}
\]

In equation 7.4: \(n_t\) represents the total patient population in a given year; \(\sigma_t\) is the actual level of ALK testing provision in a given year following a capacity investment; \(\alpha_t\) is the level of ALK testing provision at a given time in the absence of a capacity investment; \(i_t\) is the cost of the capacity investment in a given year. It should be noted that in this example the NMB is conditional on the levels of the localisation (\(\beta\)) and pathology laboratory capacity (\(\gamma\)) constraints rather than just the level of implementation. Furthermore, the level of implementation in the absence of a capacity-investment is \(\alpha_t\) as it is assumed that the commissioning awareness constraint is the only factor that limits patients’ access to ALK testing. Natural diffusion in this case study represents a natural growth in practitioners’ awareness of the commissioning awareness.

The incremental net monetary benefit (INMB) of each capacity-investment was calculated individually while keeping the other constraints in place: increased commissioning awareness (see table 7.4); centralisation of testing (see table 7.5); improved pathology laboratory staffing (see table 7.6). The total INMB of all capacity investments combined was then calculated (see table 7.7).
Table 7.5: Parameter values for capacity-investment ‘Outreach Programme’ to increase commissioning awareness

<table>
<thead>
<tr>
<th>Year</th>
<th>Patients Eligible for Testing ($n_t$)</th>
<th>Proportion of Patients Receiving Testing with no Implementation Strategy ($\alpha_t$)</th>
<th>Proportion of Patients Receiving Testing with Implementation Strategy ($\sigma_t$)</th>
<th>Proportion of Patients Receiving Localised Testing ($\beta$)</th>
<th>Level of Under-staffing in Pathology Laboratories ($\gamma$)</th>
<th>Cost of Implementation Strategy ($i_t$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2014</td>
<td>16,627</td>
<td>0.230</td>
<td>0.242</td>
<td>0.11</td>
<td>1</td>
<td>£120,080</td>
</tr>
<tr>
<td>2015</td>
<td>18,136</td>
<td>0.457</td>
<td>0.480</td>
<td>0.11</td>
<td>1</td>
<td>£0</td>
</tr>
<tr>
<td>2016</td>
<td>19,654</td>
<td>0.703</td>
<td>0.738</td>
<td>0.11</td>
<td>1</td>
<td>£0</td>
</tr>
<tr>
<td>2017</td>
<td>19,737</td>
<td>0.869</td>
<td>0.913</td>
<td>0.11</td>
<td>1</td>
<td>£0</td>
</tr>
<tr>
<td>2018</td>
<td>20,822</td>
<td>0.949</td>
<td>0.997</td>
<td>0.11</td>
<td>1</td>
<td>£0</td>
</tr>
<tr>
<td>2019</td>
<td>21,907</td>
<td>0.981</td>
<td>1</td>
<td>0.11</td>
<td>1</td>
<td>£0</td>
</tr>
</tbody>
</table>

1 From time of appraisal to current day
Table 7.6: Parameter values for a capacity-investment to centralise ALK testing

<table>
<thead>
<tr>
<th>Year</th>
<th>Patients Eligible for Testing ($n_t$)</th>
<th>Proportion of Patients Receiving Testing with no Implementation Strategy ($\alpha_t$)</th>
<th>Proportion of Patients Receiving Testing with Implementation Strategy ($\sigma_t$)</th>
<th>Proportion of Patients Receiving Localised Testing ($\beta$)</th>
<th>Level of Under-staffing in Pathology Laboratories ($\gamma$)</th>
<th>Cost of Implementation Strategy ($i_t$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2014</td>
<td>16,627</td>
<td>0.230</td>
<td>0.230</td>
<td>0</td>
<td>1</td>
<td>£173,001</td>
</tr>
<tr>
<td>2015</td>
<td>18,136</td>
<td>0.457</td>
<td>0.457</td>
<td>0</td>
<td>1</td>
<td>£0</td>
</tr>
<tr>
<td>2016</td>
<td>19,654</td>
<td>0.703</td>
<td>0.703</td>
<td>0</td>
<td>1</td>
<td>£0</td>
</tr>
<tr>
<td>2017</td>
<td>19,737</td>
<td>0.869</td>
<td>0.869</td>
<td>0</td>
<td>1</td>
<td>£0</td>
</tr>
<tr>
<td>2018</td>
<td>20,822</td>
<td>0.949</td>
<td>0.949</td>
<td>0</td>
<td>1</td>
<td>£0</td>
</tr>
<tr>
<td>2019</td>
<td>21,907</td>
<td>0.981</td>
<td>0.981</td>
<td>0</td>
<td>1</td>
<td>£0</td>
</tr>
</tbody>
</table>
Table 7.7: Parameter values for a capacity-investment to improve pathology laboratory capacity

<table>
<thead>
<tr>
<th>Year</th>
<th>Patients Eligible for Testing ($n_t$)</th>
<th>Proportion of Patients Receiving Testing with no Implementation Strategy ($\alpha_t$)</th>
<th>Proportion of Patients Receiving Testing with Implementation Strategy ($\sigma_t$)</th>
<th>Proportion of Patients Receiving Localised Testing ($\beta$)</th>
<th>Level of Under-staffing in Pathology Laboratories ($\gamma$)</th>
<th>Cost of Implementation Strategy ($i_t$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2014</td>
<td>16,627</td>
<td>0.230</td>
<td>0.230</td>
<td>0.11</td>
<td>1</td>
<td>£0</td>
</tr>
<tr>
<td>2015</td>
<td>18,136</td>
<td>0.457</td>
<td>0.457</td>
<td>0.11</td>
<td>1</td>
<td>£0</td>
</tr>
<tr>
<td>2016</td>
<td>19,654</td>
<td>0.703</td>
<td>0.703</td>
<td>0.11</td>
<td>1</td>
<td>£0</td>
</tr>
<tr>
<td>2017</td>
<td>19,737</td>
<td>0.869</td>
<td>0.869</td>
<td>0.11</td>
<td>0.42</td>
<td>£8,661</td>
</tr>
<tr>
<td>2018</td>
<td>20,822</td>
<td>0.949</td>
<td>0.949</td>
<td>0.11</td>
<td>0.42</td>
<td>£8,661</td>
</tr>
<tr>
<td>2019</td>
<td>21,907</td>
<td>0.981</td>
<td>0.981</td>
<td>0.11</td>
<td>0</td>
<td>£27,925</td>
</tr>
</tbody>
</table>
Table 7.8: Parameter values for three combined capacity-investments in the provision of ALK testing

<table>
<thead>
<tr>
<th>Year</th>
<th>Patients Eligible for Testing ((n_t))</th>
<th>Proportion of Patients Receiving Testing with no Implementation Strategy ((\alpha_t))</th>
<th>Proportion of Patients Receiving Testing with Implementation Strategy ((\sigma_t))</th>
<th>Proportion of Patients Receiving Localised Testing ((\beta))</th>
<th>Level of Under-staffing in Pathology Laboratories ((\gamma))</th>
<th>Cost of Implementation Strategy ((i_t))</th>
</tr>
</thead>
<tbody>
<tr>
<td>2014</td>
<td>16,627</td>
<td>0.230</td>
<td>0.242</td>
<td>0</td>
<td>1</td>
<td>£293,081</td>
</tr>
<tr>
<td>2015</td>
<td>18,136</td>
<td>0.457</td>
<td>0.480</td>
<td>0</td>
<td>1</td>
<td>£0</td>
</tr>
<tr>
<td>2016</td>
<td>19,654</td>
<td>0.703</td>
<td>0.738</td>
<td>0</td>
<td>1</td>
<td>£0</td>
</tr>
<tr>
<td>2017</td>
<td>19,737</td>
<td>0.869</td>
<td>0.913</td>
<td>0</td>
<td>0.42</td>
<td>£5,147</td>
</tr>
<tr>
<td>2018</td>
<td>20,822</td>
<td>0.949</td>
<td>0.997</td>
<td>0</td>
<td>0.42</td>
<td>£5,147</td>
</tr>
<tr>
<td>2019</td>
<td>21,907</td>
<td>0.981</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>£16,563</td>
</tr>
</tbody>
</table>
7.3.7 Sensitivity Analysis

Each of the capacity-investments, as defined in the base case analysis, are assumed to improve patient access to, or the cost-effectiveness of a test-treat intervention which is already cost-effective. The sensitivity analysis considered an alternative scenario in which all patients received localised testing in 2014 and the test-treat intervention was not cost-effective at the time of evaluation. This meant that in 2014, ALK testing and crizotinib, would have an ICER of £56,367 per QALY gained and a NMB of £-497,372. The value of each capacity-investment was first determined separately given this alternative starting scenario and then the combined value of all three capacity-investments was calculated.

As all ALK testing was assumed to be provided by localised laboratories in the sensitivity analysis for this study, the cost of the centralisation implementation strategy was increased to reflect the need to convince more laboratories to send their testing to other centres. The cost of validating testing in the 29 laboratories (£145,000) and the development costs of the outreach programme (£13,048) remained the same. However, as 147 laboratories needed to be visited rather than 19 this led to a variable cost of £2,198,091 and a total cost of £2,356,139.

7.4 Results

This section outlines the results of the dynamic value of implementation analysis of the value of three capacity-investments to improve the implementation of ALK testing to guide treatment with crizotinib. The value of each capacity-investment was first calculated separately and illustrated graphically. The value of all three capacity-investments combined is then calculated and illustrated. The second part of this results sections described the results of a sensitivity analysis in which the capacity-investments are valued when the intervention is originally delivered exclusively through localised laboratories and is not deemed to be cost-effective.

7.4.1 Base Case Analysis

All three of the selected individual capacity-investments yielded a positive incremental net monetary benefit indicating that they would improve the total amount of health produced in the health system by increasing the proportion of patients who had access to ALK testing. Investing in
improving awareness of ALK test commissioning resulted in an INMB of £760,702 which is equivalent to 15.21 QALYs in the additional 3,589 patients with NSCLC who received testing (see figure 7.3). The capacity-investments addressing the quality improving constraints of localisation and pathology capacity had higher INMB’s of £3,557,595 (see figure 7.4) and £3,392,875 (see figure 7.5), respectively. This equates to expected QALY gains of 71.15 and 67.86, respectively.

**Figure 7.3: The incremental monetary net benefit of a capacity-investment to improve commissioning awareness**

![Graph showing total discounted INB (£) over years](image)

Footnote: Using a discount rate of 3.5% per year

Figure 7.3 shows the impact of the capacity-investment to improve awareness of how tests are commissioned. While the implementation strategy yields a positive INMB, its impact is relatively small given that there is a natural growth in the awareness of commissioning arrangements without the capacity-investment.
Figure 7.4: The incremental net monetary benefit of a capacity-investment to promote centralising *ALK* testing

Figure 7.4 shows the impact of strategy capacity-investment to centralise *ALK* testing which in turn improves the quality of testing. This capacity-investment has a bigger impact than raising commissioning awareness as it improves the cost-effectiveness of the intervention for all patients receiving *ALK* testing. This finding is exhibited by the widening gap between the actual implementation and current implementation curves as more patients receive testing through natural diffusion.
The impact of a capacity-investment (figure 7.5) to improve pathology laboratory capacity was large but there was a significant delay until this impact is seen. In 2017 the additional biomedical scientists entered the workplace resulting in a rise in the NMB. In 2019 the consultant histopathologists also entered the workplace and raised NMB to nearly the level of perfect implementation. The NMB did not reach the level of perfect implementation as 11% of patients are receiving local IHC testing which is slightly more expensive than if testing were fully localised.
Figure 7.6: The incremental net monetary benefit of three capacity-investments to improve the provision of ALK testing.

When all three of the capacity-investments are combined the resulting INMB is £6,247,486 or 124.95 QALYs. This result was slightly lower than the sum of the INMB all the individual capacity investments due to interactions between the capacity constraints (see section 6.3). Figure 7.6 shows that when all three capacity-investments are made, the use of ALK testing reaches perfect implementation and its maximum societal value. Despite the impact of the commissioning awareness implementation strategy in improving the implementation of ALK testing, it takes five years from approval to be available for all patients in its most cost-effective form. This is because the commissioning awareness investment only raises ALK test availability by 5% compared with natural diffusion and because of the time delays in hiring staff to fill vacancies and improve the quality of testing.

7.4.2 Sensitivity Analysis

The sensitivity analysis considered a scenario in which all patients received localised testing in 2014 and ALK testing and crizotinib was not cost-effective at the time of appraisal. This
assumption had a large impact on the estimated INMBs of the capacity-investments. The impact of raising commissioning awareness (figure 7.7), centralising testing (figure 7.8), improving staffing levels in pathology laboratories (figure 7.9), and combining all three capacity-investments is shown in figure 7.10.

**Figure 7.7:** The incremental net monetary benefit of investing in commissioning awareness when ALK testing and crizotinib is not cost-effective

When ALK testing and crizotinib was not initially cost-effective when appraised in 2014, raising awareness of commissioning arrangements resulted in a negative INMB of £-524,033. This finding was because implementing ALK testing and crizotinib resulted in a net societal loss and the capacity-investment increased the number of patients receiving the test-treat intervention. This was equivalent to causing a net societal loss of 10.48 QALYs.
Contrary to the reduction in the INMB of the commissioning awareness capacity-investment, the value of a capacity-investment to increase centralisation was significantly increased if all testing was initially localised. The INMB of the strategy of £31,558,366 is actually nearly ten times greater than the INMB of the strategy when most testing is centralised and ALK testing and crizotinib is cost-effective. This result is because when all ALK testing is localised, the capacity-investment is required to make the intervention cost-effective. In the first year, the capacity-investment causes a net societal loss due to its significant upfront cost. This loss is larger than that which would be experienced if no capacity-investment was undertaken to implement the ALK testing. However, this initial capacity-investment results in significant societal gains over time which rise with the natural diffusion of the ALK testing.
Figure 7.9: The incremental net monetary benefit of investing in pathology capacity when *ALK* testing and crizotinib is not cost-effective

The value of the capacity-investment to improve pathology laboratory capacity was also significantly higher when all *ALK* testing was originally localised. Improving pathology laboratory capacity makes *ALK* testing and crizotinib cost-effective and results in a large INMB of £17,190,751. The observed kinks in the diffusion curve were caused by the length of time it takes to train the new staff members required to improve capacity.
When all of three of the capacity-investments are implemented, an INMB of £36,431,395 was calculated. This estimate is equivalent to a net gain of 729 QALYs. The initial value of investment (see figure 7.10) means that there is a negative INMB in the first year but then a rapid rise in the value of capacity-investment until perfect implementation is achieved in 2019.

7.5 Discussion

This study used the dynamic value of implementation methods developed in chapter 3 to quantify the value of using capacity-investments to reduce the impact of three capacity constraints on the cost-effectiveness and net societal benefit of ALK testing and crizotinib. The VOImp methods allow for variable marginal net benefits and costs (see also Chapter 3). The VOImp methods were applied to three hypothetical capacity-investments to provide ALK testing to guide treatment with crizotinib. The capacity-investments in improving awareness of commissioning arrangements for ALK testing, centralising ALK testing, and improving the staffing levels of pathology laboratories
all yielded positive INMB indicating that investing in capacity could produce gains in the total health produced by the health system.

Despite the evidence in the static analysis that suggested that a lack of commissioning awareness was the most significant capacity constraint (see Chapter 6), this study suggested that capacity-investments to reduce the impact of localised testing or insufficient pathology laboratory staffing generated relatively more added value. The reduced impact of improving awareness of ALK test commissioning was partly because the natural diffusion of awareness of commissioning arrangements for ALK testing in this study reduced the potential value of a strategy to actively improve such awareness. Furthermore, improvements in the “quality limiting” constraints of localised testing or pathology staffing either reduced the marginal cost of testing per patient or increased the marginal benefit, improving the cost-effectiveness of ALK testing and crizotinib.

The sensitivity analysis conducted in this study provided supportive evidence that capacity-investments in improving the quality of a test-treat intervention should be prioritised over improving access to the intervention. When ALK testing and crizotinib was not cost-effective at the time of evaluation, investing in removing “access limiting” constraints meant that more patients would potentially receive a cost-ineffective intervention and there was an increasing net societal loss. Investing in centralised testing or increased pathology laboratory capacity made ALK testing and crizotinib cost-effective for all patients turning the net societal loss into a net societal gain.

Previous value of implementation studies have focussed on implementation strategies which increase the number of patients receiving a cost-effective intervention (164,165). As such, in the terminology outlined in this thesis, they represent investments in ‘access limiting’ constraints. For example, the study by Mewes et al., (2017) which was used to inform the capacity-investments used in this chapter aimed to determine the value of implementation strategies to increase the adherence of health professionals to guidelines on physical exercise for cancer survivors (165). In this case the non-adherence of health professionals to these guidelines can be seen as an “access limiting” capacity constraint with the exercise guideline as a cost-effective intervention. Similarly, Whyte et al., (2016) estimated the value of implementation for increasing uptake of natriuretic peptide testing in patients with suspected heart failure (164). This study represented an investment in improving access to a cost-effective intervention. No value of implementation studies have previously investigated the value of implementation strategies which improve the cost-effectiveness of the intervention rather than solely improving patient access.
Limitations

There were some limitations to this dynamic value of implementation study. The capacity-investments outlined in this study were hypothetical, by necessity, and their parameterisation was based on a number of assumptions. This was necessary due to the availability of the required data to populate the model.

To provide a pragmatic illustration of the method to evaluate the value of capacity-investments to improve the implementation of ALK testing and crizotinib of precision medicine, probabilistic sensitivity analysis was not conducted. In future developments of this method, it will be important to incorporate parameter uncertainty into the value of implementation approach.

An additional key area of uncertainty which was not considered in this paper was in the costs and effects of the capacity-investments. Even if the cost-effectiveness of the intervention was known with certainty, this variation in the effects and costs of the capacity investments would mean that sometimes it would not be valuable to invest in capacity. When the uncertainty in the capacity investments and uncertainty in the cost-effectiveness of the intervention is combined there will likely be a multiplicative effect resulting in wide variations in the estimated INMB of the capacity investments. Value of information analysis may therefore be useful in determining whether investment in capacity investments or further research to reduce uncertainty should be prioritised (169,170). As was suggested in chapter 2, this uncertainty can itself be seen as a capacity constraint to introducing new examples of precision medicine (161).

7.6 Conclusion

This study has shown how dynamic value of implementation methods can be used to estimate the value of investing in improving the capacity of the health system (capacity-investments) to deliver an example of precision medicine (ALK testing and crizotinib). Capacity-investments were shown to improve the relative cost-effectiveness of ALK testing and crizotinib. No previous value of implementation studies have investigated the value of investments to address constraints which impact the cost-effectiveness of an intervention and this is a promising area of future research.
7.7 Summary of Chapter 7

Chapter 7 demonstrated how the dynamic value of implementation method developed in chapter 3, which allows for varying marginal costs, can be used to determine the value of investing in improving the capacity of the health system to implement an example of precision medicine.

Chapter 8 will discuss the collated findings from the six empirical chapters in terms of the proposed method for accounting for capacity constraints, the limitations of the methods, and the implications of the findings of these studies for policy making and research.
Chapter 8

Discussion and Conclusions

8.1 Chapter 8 Summary

The overall aim of this PhD was to identify and quantify the impact of including capacity constraints in decision-analytic model-based cost-effectiveness analysis to better inform resource allocation decisions and introduction of precision medicine into clinical practice. Chapter 8 provides a summary of the key findings from the six empirical studies in this thesis. Chapter 8 also discusses the overall contribution of the six empirical studies in terms of the proposed method for accounting for capacity constraints, the limitations of the methods, and the implications of the findings of these studies for policy making and research.

8.2 Contribution of this Thesis to Knowledge

This thesis made a number of novel contributions to knowledge.

1. A definition of health system capacity constraints was proposed in chapter 1 and this was subsequently published in the systematic review reported in chapter 2.
2. The systematic review reported in chapter 2 demonstrated that current economic evaluations of precision medicine rarely account for capacity constraints but highlighted which methods had been used in the few studies that had sought to quantify the impact of such constraints.
3. In chapter 3, it was demonstrated how varying marginal costs and benefits could be incorporated into value of implementation analysis and how these methods could be used to quantify the impact of varying levels of implementation on the total net benefit produced by an example of precision medicine.
4. In chapter 4 a typology of barriers to the implementation of examples of precision medicine for NSCLC was developed. This may aid in the identification of barriers to implementation for future examples of precision medicine.
5. Chapter 6 demonstrated how the static value of implementation methods developed in chapter 3 could be used to quantify the impact of a number of capacity constraints to an example of precision medicine. An additional finding was that there may be two types of
capacity constraint: “access limiting” constraints which prevent patients from receiving the intervention but do not affect its cost-effectiveness; and “quality limiting” constraints which affect the cost-effectiveness and total net monetary benefit of the intervention.

6. Chapter 7 demonstrated how the dynamic value of implementation methods developed in chapter 3 can be used to determine the value of investments to reduce the impact of capacity constraints to an example of precision medicine.

7. Combined, chapters 4, 5, 6, and 7 demonstrate an approach to identifying capacity constraints, incorporating these in a decision analytic model, quantifying the impact of these constraints on the cost-effectiveness and net monetary benefit of an intervention, and quantifying the value of investments aimed at overcoming the constraints.

8.3 Summary of Key Findings

This section summarises the key findings of the thesis. The first research question of this thesis was “have published decision-analytic model-based cost-effectiveness analyses of examples of precision medicine accounted for health system capacity?” This question was addressed in chapter 2 using a meta-review of systematic reviews of economic evaluations of examples of precision medicine which identified a sample of 222 primary economic evaluations. Under a sixth (33) of these studies discussed potential capacity constraints to the introduction of examples of precision medicine with the constraints falling into four categories: budget constraints; quality of the testing process; ease of test use in clinical practice; and the need for economic evidence to reduce decision uncertainty. However, despite this broad range of potential barriers to examples of precision medicine, only nine of the included studies used methods which could quantitatively account for the capacity constraints. This suggested there was a large gap in the literature base with regards to evaluating the impact of such capacity constraints on the cost-effectiveness of examples of precision medicine.

The second research question addressed by this thesis was “how could health system capacity constraints be included decision-analytic model-based cost-effectiveness analysis of an example of precision medicine”. The nine studies identified in the meta-review in chapter 2 which had incorporated capacity constraints used a range of methods to do so. These were divided into static methods for accounting for capacity constraints and dynamic methods for accounting for capacity constraints. Static methods were defined in chapter 2 as methods which “produce a single cost-effectiveness estimate which takes account of imperfect implementation for one cohort of patients”
In studies which quantified the static impact of capacity constraints, incorporating these constraints had an impact on the cost-effectiveness of the examples of precision medicine in some economic evaluations (153,155,293) and in one example changed the ordering of the most cost-effective strategy (139). It is therefore possible that economic evaluations of examples of precision medicine which omit capacity constraints may provide a misleading representation of the cost-effectiveness of the intervention.

Dynamic methods for accounting for capacity constraints were defined in chapter 2 as “methods that account for capacity constraints which allow the impact of barriers or constraints to change over time and/or in multiple patient cohorts” (161). Dynamic methods for accounting for capacity constraints have an advantage over static methods in that they move beyond a simple yes or no decision as to whether an example of precision medicine should be approved to allowing questions such as when might the intervention be cost-effective and under what circumstance. This was demonstrated by Retèl et al., (2012) who identified a number of barriers to the introduction of a 70-gene test to predict the risk of breast cancer recurrence (136). While this study focussed on the natural change in capacity over time, it was proposed in chapter 2 that a method which could value active investments in improving capacity with the aim of improving the cost-effectiveness of an intervention and access to the intervention for patients would be a valuable contribution.

The key requirements derived from chapter 2 for a method which could account for capacity constraints were: that the method could account for the impact of the constraints on the costs and benefits of the intervention; that the method could quantify the impact of capacity constraints on the cost-effectiveness of the intervention at different time points; and that the method could quantify the value of investments to improve health system capacity with the aim of improving the cost-effectiveness of the example of precision medicine.

Chapter 3 reported the development of static and dynamic value of implementation methods which aimed to address these requirements. This involved adapting existing methods to allow for the varying marginal costs and benefits which may arise due to capacity constraints. The study by Retèl et al., (2012) identified in chapter 2 was used as a case study to illustrate the consequences of varying marginal costs and benefits for the total societal value of an intervention at different levels of implementation arising due to capacity constraints. It was shown that the total societal value of the intervention depended on how it was implemented and for how many people. A significant finding was that even though increasing implementation improved the incremental cost-
effectiveness ratio of the 70-gene recurrence score assay, this yielded increasing losses to total societal net benefit at low levels of implementation as more participants received an intervention that was not cost-effective. This suggested that the ICER alone as an output from a CEA may not capture the full impact of capacity constraints and that the measures of total societal net benefit used in value of implementation analysis may be more appropriate.

The second stage of this thesis aimed to demonstrate how the methods developed in chapter 3 could be applied to an example of precision medicine. The next research question addressed in the thesis was: “what are the health system capacity constraints which may impede the introduction of examples of precision medicine into clinical practice?” To answer this question, a qualitative interview study was undertaken which aimed to produce a typology of barriers to the implementation of examples of precision medicine in the field of non-small cell lung cancer. In total, 17 barriers which had been faced in introducing examples of precision medicine for NSCLC were identified and these fell into five themes: the managed entry of precision medicine for NSCLC; the commissioning and reimbursement of precision medicine for NSCLC and specifically the test component of precision medicine; the complexity of the logistics around providing tests; opinions about whether test provision should be localised or centralised; and opinions about future developments, including potential barriers to their introduction, in precision medicine for NSCLC. A key finding of this study was that the majority of the barriers to introducing examples of precision medicine for NSCLC lay in providing rapid, high quality, well-funded testing while few lay in providing the drug component. Many of these barriers bore similarities to those identified in the systematic review in chapter 2, suggesting that they may be common across different disease areas.

Chapters 5 and 6 sought to address the research question: “how can the impact of health system capacity constraints be measured in decision-analytic model-based cost-effectiveness analyses of an example of precision medicine?” First, a decision-analytic model was constructed to estimate the cost-effectiveness of an example of precision medicine, ALK testing to guide crizotinib treatment, in the absence of capacity constraints. This model was a replication of the decision-analytic model submitted by the manufacturer of crizotinib and as such represents current practice with regards to providing an estimate of the cost-effectiveness of an example of precision medicine. While reproducing the model proved challenging due to the presence of redacted parameter information, the baseline model provided similar results to the source model.
In chapter 6, the baseline model was adapted to include three key capacity constraints limiting the introduction of *ALK* testing to guide crizotinib which had been identified in chapter 4: lack of awareness about how *ALK* testing was commissioned; degree of centralisation of immunohistochemistry testing; and ability to conduct *ALK* testing dependent on whether centralised or localised services were used. The baseline decision-analytic model developed in chapter 5 was adapted to incorporate these constraints and new parameter values found to represent their impact on the provision of *ALK* testing. The static value of implementation methods developed in chapter 3 which allow for varying marginal costs and benefits were used to quantify the impact of the constraints on the cost-effectiveness and total net monetary benefit of the intervention. Lack of awareness about how *ALK* testing was commissioned had the biggest impact on the total net benefit of the intervention despite not having an impact on the ICER of the intervention. Conversely the degree of centralisation of IHC testing and ability to conduct *ALK* testing constraints reduced the ICER of the intervention but only had a limited impact on the total net monetary benefit.

The results of chapter 6 had a number of key implications. Capacity constraints may impact the results of a decision-analytic model based cost-effectiveness analysis in two ways: they may reduce the number of patients who have access to an intervention; or they may impact on the marginal costs or benefits of the intervention and therefore its cost-effectiveness. In this thesis capacity constraints with these characteristics have been labelled “access limiting” or “quality limiting” constraints respectively. This finding itself has implications for reporting the results of an economic evaluation that includes capacity constraints. As was suggested in chapter 3, the ICER as an output is incapable of capturing the full impact of capacity constraints in an economic evaluation of an example of precision medicine. This limitation is because the ICER does not present any information about changes to the number of patients who can receive an example of precision medicine which may arise due to “access limiting” capacity constraints. To fully account for the impact of capacity constraints, a measure which incorporates both the ratio of costs to benefits and the population size should be used. Such outputs can be represented using total net monetary or health benefits used in value of implementation analysis.

The results of chapter 6 suggested that while *ALK* testing and targeted treatment with crizotinib would have been cost-effective in 2014 despite the presence of capacity constraints, the total net benefit produced for the health system would have fallen far short of the potential for the intervention in the absence of the constraints. This result in practice means that a significant
number of patients who potentially could have benefitted from crizotinib would not have received it due to the presence of capacity constraints. The final research question in this thesis was “can an economic perspective be used to evaluate the impact of the removal of health system capacity constraints to the introduction of an example of precision medicine?” This research question was addressed in chapter 7 that reported a study which used the dynamic value of implementation methods developed in chapter 3 to determine the value of three hypothetical investments in health system capacity. Each of the three investments represented a cost-effective use of resources but raising awareness of ALK test commissioning was the least valuable. While this appears contradictory given that the lack of awareness of commissioning arrangements was found to be the biggest constraint in chapter 6, the value of investing in improving awareness is eroded over time by the natural diffusion of the intervention in practice. Conversely, while the degree of centralisation of ALK testing and ability to offer ALK testing constraints had little impact in the static model, investments to remove these constraints were significantly more valuable.

Chapter 7 also explored the value of capacity investments when the example of precision medicine was not cost-effective due to the presence of capacity constraints. In this case, reducing the impact of “access limiting” constraints such as awareness of ALK test commissioning arrangements resulted in a large net monetary loss as more patients received an intervention which was not cost-effective. However, investments in addressing the “quality limiting” constraints and making the intervention cost-effective were extremely valuable. This suggests that the priority for decision makers faced with the decision as to whether to adopt an example of precision medicine is addressing any capacity constraints which when present result in the intervention not being cost-effective. When these constraints have been addressed and the intervention does represent a cost-effective use of resources, investments targeting “access limiting” constraints can be applied to allow more patients to receive the intervention.

8.4 Summary of the Methodological Contribution of this Thesis

This thesis has used a retrospective case study of an example of precision medicine to show how capacity constraints can be identified and incorporated into an economic evaluation and how strategies for overcoming capacity constraints can be evaluated. By addressing the aims and objectives of each chapter in this thesis, a method was developed for including capacity constraints
in economic evaluation of precision medicine and for valuing capacity investments. This approach comprises the following steps:

1. Define the decision problem using the PICO (population, intervention, comparator, outcome) acronym as a guide. This decision problem should reflect the presence of capacity constraints which may impact on the participants who can receive the example of precision medicine, how the intervention is offered to participants, how the comparator is offered to participants, and the outcomes produced. The outcomes should include total net benefit which has been shown in this thesis to be necessary to capture all of the potential impacts of capacity constraints on the health produced for patients by the health care system.

2. Identify potential capacity constraints to the introduction of an example of precision medicine using qualitative methods with stakeholders in the implementation process.

3. Develop a decision analytic model to evaluate the example of precision medicine using a conventional approach to economic evaluations but including total net benefit as an output alongside the ICER. These outputs present the cost-effectiveness and value in terms of total net benefit of the intervention under a scenario of perfect implementation.

4. Identify the capacity constraints that are likely to have a significant impact on the cost-effectiveness of the intervention of the total number of patients who receive the intervention.

5. Incorporate the selected capacity constraints into the decision analytic model by changing its structure and input parameters to reflect the ability of the health system to provide the intervention at the time of potential approval. The static impact of the constraints is measured by the value of perfect implementation which is the difference in the net benefit achieved when the constraints are in place (the value of current implementation) and the net benefit accrued under perfect implementation.

6. Identify potential strategies and investments to reduce the impact of the capacity constraints. Predict the potential benefits of these strategies in terms the degree to which they address the constraints and determine their likely cost. Convert the decision analytic model into a dynamic, multiple cohort model and incorporate changes in the size of the patient population and the natural diffusion of the intervention. Determine the value of the strategies to overcome capacity constraints by calculating the value of actual implementation which is the difference in total net benefit when the implementation strategies are enacted and when the intervention is left to naturally diffuse.
8.5 Limitations

The limitations of each component of the proposed method for accounting for capacity constraints in economic evaluations of precision medicine are detailed in each chapter. This section outlines some broader limitations which apply to the use of the method as a whole. In summary, these limitations are: difficulties in applying the method to a prospective example; balancing model complexity and clarity; and the utility of the method to decision makers. Furthermore, within the general discussion of the difficulty in applying the method to a prospective economic evaluation, three specific but linked issues are explored: the need to forecast potential barriers and parameter values; the lack of data on the impact of potential barriers; and the need for additional qualitative research to support the application of the method.

8.5.1 Applying the Proposed Method to a Prospective Economic Evaluation

This study used a retrospective example of an economic evaluation of an existing application of precision medicine to illustrate how the presence of health system capacity constraints could be quantified. However, in practice the proposed method to quantify the impact of capacity constraints is likely to be used in prospective evaluations, for example in early economic evaluations of health technologies or alongside technology appraisals which take place before an intervention is widely used in the healthcare service. This means that the researchers conducting the economic evaluation will have to predict the potential barriers to the introduction of an example of precision medicine.

Forecasting these barriers and their potential impact is likely to be difficult. In the qualitative interviews conducted in chapter 4, participants were good at identifying issues that were faced in introducing existing examples of precision medicine but identified fewer barriers to future innovations. To aid in the identification of barriers to the implementation of a prospective example of precision medicine, different and more in-depth qualitative methods may be required.

As was noted in section 4.5, a potential source of these more complex qualitative methods for identifying the barriers that will be faced in introducing an example of precision medicine is the field of implementation science. Researchers working in this field have developed a number of theories and frameworks to aid in identifying the key barriers and facilitators to implementing an intervention. The constructs included in these frameworks are broad, linking with the broad definition of capacity constraints taken in this thesis. For example, the Consolidated Framework for Implementation Research includes constructs such as the strength of evidence for the intervention,
the impact of external policies, how connected an organisation is to other organisations, organisational culture, and individual attitudes to the intervention and the organisation (15). The use of such a framework may be useful in developing an interview schedule to help identify the key barriers to implementation.

The integration of implementation science methods will likely result in a more realistic representation of the potential challenges of implementing complex interventions such as approaches to precision medicine. Including all of these barriers in an economic model may be difficult, and probably not feasible, and as such the research team may have to condense these into a number of key barriers to be modelled. The identification of these key barriers should involve stakeholders in the process and again involve the use of more robust qualitative methods. Potential examples of these methods include the scenario drafting and Delphi methods used by Retèl et al., (2012) (24). Using Delphi methods involves collating the views of key stakeholders in order to understand the degree of consensus on likely scenarios which describe the potential experiences that stakeholders may face in introducing the intervention. By reducing the number of key barriers, focussing on those in which there is consensus about the relevance, these can then be readily incorporated into decision analytic models that aim to quantify their impact on the cost-effectiveness and total net monetary benefit of the intervention.

An additional problem with forecasting the potential impact of capacity constraints on a future example of precision medicine will be faced in identifying quantitative parameters. Even in the retrospective example presented in this thesis, it was difficult to identify parameters representing the proportion of patients with access to testing or the impact of understaffing in pathology laboratories. To provide estimates of these parameters, researchers may be required to use expert elicitation to explore stakeholders’ perceptions of the future potential impact of capacity constraints (291,294).

8.5.2 Balancing Model Complexity and Clarity

In this thesis, a relatively simple approach to incorporating capacity constraints into a decision analytic model has been used. This approach, reported in chapter 6, was to add branches to a decision tree and identify new parameters for the probabilities attached to these branches and for the costs and health utilities associated with patients experiencing these altered testing pathways. More complex methods of incorporating capacity constraints into decision-analytic models are available including individual level simulation models (48). These more detailed models may more
accurately reflect the mechanisms through which capacity constraints affect costs and outcomes for patients. For example, they can model the impact of patients having to queue of services due to constraints (295). However, while the case study used in this thesis did not use simulation methods to model the effect of constraints, such methods could be integrated within the method proposed in this thesis for quantifying the impact of the constraints using static and dynamic value of implementation methods.

8.5.3 Utility of the Results to Decision Makers

The methods developed in this PhD can help to provide evidence as to which capacity constraints may have the most significant impact on the cost-effectiveness and total net monetary benefit of an example of precision medicine. They can also provide evidence as to which investments in improving health system capacity may be valuable and which should be prioritised. However, to date the utility of these findings for decision makers has not been explored.

There are two potential problems which may limit the utility of these findings for decision makers. In this thesis, the methods have been applied to a case study based on a technology appraisal conducted by NICE. Investments to improve the capacity of the health system to provide an example of precision medicine may have to be initiated by a range of stakeholders and budgets. For example, when considering the hypothetical investments outlined in chapter 7, a national campaign to improve awareness of commissioning arrangements for ALK testing may have to be initiated by a body like Public Health England, while improving the staffing of pathology laboratories may be the responsibility of Health Education England (296). Co-ordinating the efforts of different bodies in improving the provision of a specific intervention may prove difficult.

In addition to attempting to co-ordinate the actions of multiple stakeholders to implement a specific example of precision medicine, these stakeholders will also each have their own budgets from which any capacity investments will be funded. These budgets must also be used for a number of other strategies and interventions and so there is an opportunity cost of investing in reducing capacity constraints for a specific intervention. Furthermore, these opportunity costs may be different for different stakeholders. In addition, stakeholders may have different goals for their budgets and these may not necessarily be the maximisation of QALYs.
8.6 Implications of this Thesis

This section outlines the potential implications of the findings of this thesis for health policy and suggests some potential avenues for future research.

8.6.1 Implications for Health Policy

A key finding of this thesis is that health system capacity constraints can reduce the cost-effectiveness and net monetary benefit of examples of precision medicine. However, as was reported in chapter 2, economic evaluations of examples of precision medicine rarely account for such capacity constraints. The current approach to conducting economic evaluations has been argued to provide an estimate of the long run cost-effectiveness on an intervention while ignoring the short run (77,109,111). When capacity constraints to the introduction of an intervention exists in the short run, such estimates may not convey all of the information needed by policy makers to inform the adoption and implementation of such interventions.

The discrepancy between the short run and long run cost-effectiveness of interventions has significant implications for agencies involved in healthcare technology assessment such as NICE or CADTH. It is possible that interventions may appear cost-effective based on estimates from traditional decision analytic models but in reality will not be cost-effective in the short term until investments are made in improving health system capacity. If the approval of the technology is based on estimates of the long run cost-effectiveness of the intervention, then the burden of such investments will fall on the health system. This is particularly true if pharmaceutical companies price their drugs such that their estimated cost-effectiveness falls just below the threshold recommended by an HTA agency. In this case, incorporating the costs of investing in capacity into a health technology appraisal would likely reduce the ceiling price of the drug. Omitting the short term capacity investments may therefore mean that the health system must cover the cost of investing in capacity and will also pay more than is necessary for the drug component of the precision medicine.

Even when capacity constraints do not impact the cost-effectiveness of an intervention, they may restrict patient access to beneficial interventions, thereby reducing the net monetary benefit produced by the health system. The use of total net monetary benefit as an output, as is used in value of implementation analysis, is rare in economic evaluations. For example, while net benefit is mentioned in the glossary of the NICE guide to the methods of technology appraisal, the document
almost exclusively focuses on using ICERs as the outcome of submitted economic evaluations (44). However, the ICER only confers information about the impact of capacity constraints which affect the marginal incremental costs and benefits of the intervention. In order to quantify the impact of capacity constraints in terms of patients missing out on tests and treatments which may improve their quality and length of life, total population net benefit measures such as the current, actual and perfect value of implementation should be more widely adopted in health economics.

This thesis has used the area of precision medicine as a case study due to the inherent complexity in implementing such interventions. However, capacity constraints may be a significant issue for a wide range of health care interventions and policies beyond precision medicine. For example, upfront investment may be needed to train surgeons in a new technique and learning effects may impede the health outcomes for patients while the surgeons gain experience. In another example, a new strategy to identify patients with cancer through a change in the screening programme may be hindered by a lack of trained staff such as radiographers. The approach to accounting for health system capacity constraints described in this thesis can, in theory, be applied to any area of medicine where such capacity constraints may impact on the cost-effectiveness of the intervention or the number of patients receiving the intervention.

8.6.2 Implications for Research

Precision medicine and the problems associated in conducting economic evaluations of such interventions continues to be an area of expanding research in health economics (97,98,161). This thesis proposed a method for addressing one such problem: incorporating health system capacity constraints. However, the research conducted for this thesis has also raised the potential for additional lines of enquiry which are now summarised:

1. To apply the proposed method to an example of precision medicine under development. A potential application for the proposed methods is in the early economic evaluation of interventions which are still being developed. This is proposed because the prospective use of the proposed methods may help researchers and clinicians to anticipate future barriers to the introduction of examples of precision medicine and to enact strategies to smooth the implementation of the intervention.
2. In order to apply the proposed method to a prospective example, it may be necessary to use expert elicitation to identify potential parameter values to reflect the impact of future constraints. This will be difficult as participants would have to predict the future value of a parameter which may be very uncertain.

3. In chapter 3, value of implementation methods were adapted to account for varying marginal costs and benefits and these were applied to a case study in chapter 6 and 7. However, a key assumption was made in that the cost-effectiveness of the intervention was known to the decision maker with certainty. In practice, the decision maker would have a choice of whether to adopt and implement and intervention based on the current evidence or to conduct more research to remove some of this uncertainty. It is possible to use value of implementation and information approaches to aid in this decision making but further research on the impact of varying marginal costs and benefits on this choice is needed (110,165). Incorporating value of implementation methods with the value of implementation methods outlined in this thesis will provide evidence to decision makers as to whether it is better to invest in implementing an intervention with varying marginal costs or benefits or in undertaking more research. This may be particularly important as it was demonstrated in chapter 2 that a lack of evidence as to the clinical and economic benefits of an intervention may itself act as a capacity constraint to implementation.

4. Chapter 4 used qualitative interviews to identify potential barriers to the introduction of examples of precision medicine for NSCLC. The structured framework method used in this study was relatively simple and may not be as useful for identifying barriers to future examples of precision medicine. Exploring the utility of alternative approaches to identifying capacity constraints, such as developing a theory of change or using a scenario drafting approach, would be an interesting line of enquiry.

5. The utility of the proposed method for accounting for capacity constraints and the results it produces for decision makers has not been explored in this thesis. It would be useful to receive input on the degree to which the information provided meets the needs of decision makers in quantifying the impact of constraints and prioritising capacity investments.

8.7 Conclusion

As complex interventions involving both a testing and treatment component, examples of precision medicine face a range of barriers to their implementation. This thesis has demonstrated that while
these capacity constraints are commonly ignored in economic evaluations, they may have a significant impact on the cost-effectiveness and societal net benefit of these interventions at the time of appraisal and over the course of their implementation.

This thesis proposed a method to identify capacity constraints to the introduction of examples of precision medicine, incorporate these into an economic model, and quantify their impact using a static value of implementation analysis. A dynamic value of implementation analysis, which allowed for the varying marginal costs and benefits which may result from capacity constraints, was then used to demonstrate how investments in capacity to improve the implementation of an example of precision medicine could be evaluated.
9 References


34. Smith RD, Sach TH. Contingent valuation: What needs to be done? [Internet]. Vol.


46. Glick HA, Doshi JA, Sonnad SS, Polsky D. Economic Evaluation in Clinical Trials.


108. Sheldon TA. What’s the evidence that NICE guidance has been implemented? Results from a national evaluation using time series analysis, audit of patients’ notes, and interviews. Br Med J [Internet]. 2004;329(7473):999–0. Available from: http://www.bmj.com/cgi/doi/10.1136/bmj.329.7473.999


Available from:


http://dx.doi.org/10.1016/j.ejogrb.2012.03.025


158. Guidelines for the Economic Evaluation of Health Technologies: Canada 4th


168. Willan AR, Eckermann S. Optimal clinical trial design using value of information methods with imperfect implementation. Health Econ [Internet]. 2009 May 1 [cited


204. 1st Class Secretarial Services. 1st Class Secretarial [Internet]. 2019. Available from: https://www.1stclass.uk.com/


224. R Core Team. R: A language and environment for statistical computing [Internet]. Vienna, Austria; 2019. Available from: https://www.r-project.org


250. Bornmann I. DigitizeIt [Internet]. 2019. Available from: https://www.digitizeit.de/contact.html#Contact

251. Ternau T. A Package for Survival Analysis in S_ [Internet]. 2015. Available from:


255. Latimer N. NICE DSU TECHNICAL SUPPORT DOCUMENT 14: SURVIVAL ANALYSIS FOR ECONOMIC EVALUATIONS ALONGSIDE CLINICAL TRIALS -EXTRAPOLATION WITH PATIENT-LEVEL DATA REPORT BY THE DECISION SUPPORT UNIT. 2011.


262. McManus E, Turner D, Gray E, Khawar H, Okoli T, Sach T. Barriers and Facilitators to Model Replication Within Health Economics. Value Heal [Internet].


Local opinion leaders: effects on professional practice and health care outcomes.


Appendices
Appendix 2.1: Wright SJ, Newman WG, Payne K. Accounting for Capacity
Constraints in Economic Evaluations of Precision Medicine: A Systematic Review.
Pharmacoeconomics. 201937(8):1011–27¹

¹Title formatting has been changed to avoid clashes with the formatting of the thesis
Accounting for Capacity Constraints in Economic Evaluations of Precision Medicine: A Systematic Review

Stuart J. Wright1 · William G. Newman2,3 · Katherine Payne1

Published online: 13 May 2019
© The Author(s) 2019

Abstract

**Background and Objective** Precision (stratified or personalised) medicine is underpinned by the premise that it is feasible to identify known heterogeneity using a specific test or algorithm in patient populations and to use this information to guide patient care to improve health and well-being. This study aimed to understand if, and how, previous economic evaluations of precision medicine had taken account of the impact of capacity constraints.

**Methods** A meta-review was conducted of published systematic reviews of economic evaluations of precision medicine (test–treat interventions) and individual studies included in these reviews. Due to the volume of studies identified, a sample of papers published from 2007 to 2015 was collated. A narrative analysis identified whether potential capacity constraints were discussed qualitatively in the studies and, if relevant, which quantitative methods were used to account for capacity constraints.

**Results** A total of 45 systematic reviews of economic evaluations of precision medicine were identified, from which 222 studies focusing on test–treat interventions, published between 2007 and 2015, were extracted. Of these studies, 33 (15%) qualitatively discussed the potential impact of capacity constraints, including budget constraints; quality of tests and the testing process; ease of use of tests in clinical practice; and decision uncertainty. Quantitative methods (nine studies) to account for capacity constraints included static methods such as capturing inefficiencies in trials or models and sensitivity analysis around model parameters; and dynamic methods, which allow the impact of capacity constraints on cost effectiveness to change over time.

**Conclusions** Understanding the cost effectiveness of precision medicine is necessary, but not sufficient, evidence for its successful implementation. There are currently few examples of evaluations that have quantified the impact of capacity constraints, which suggests an area of focus for future research.

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s40273-019-00801-9) contains supplementary material, which is available to authorized users.

Stuart J. Wright
stuart.wright-2@manchester.ac.uk

Key Points for Decision Makers

Examples of precision medicine are complex interventions and limited health system capacity may impede their adoption into clinical practice.

Capacity constraints may have an impact on the cost effectiveness of examples of precision medicines and should be included in economic evaluations of such interventions.

Evidence as to the value of removing capacity constraints over time in terms of improving the cost effectiveness of examples of precision medicine may be also be useful to decision makers in guiding strategies to improve...
1 Manchester Centre for Health Economics, Division of Population Health, Health Services Research and Primary Care, The University of Manchester, Oxford Road, Manchester M13 9PL, UK

2 Manchester Centre for Genomic Medicine, Division of Evolution and Genomic Sciences, The University of Manchester, Manchester, UK

3 North West Genomic Laboratory Hub, Manchester Centre for Genomic Medicine, Manchester University NHS Foundation Trust, Manchester, UK
Introduction

Precision (stratified or personalised) medicine is underpinned by the premise that it is feasible to identify known heterogeneity using a specific test or algorithm in patient populations to guide patient care to improve health and well-being [1]. There is no unified definition of precision (stratified or personalised) medicine, but in practice to date it refers to using some test-treat combination to target an intervention [1, 2]. A variety of mechanisms are under- going development to identify such heterogeneity in the outcomes of interventions and progression of disease in populations of patients informed by genomic, proteomic, transcriptomic and metabolomic strategies [2]. The ability to determine which patients might be more likely to benefit from a treatment, avoid adverse effects or experience more severe disease has driven the theoretical arguments that precision medicine is a cost-effective use of healthcare resources [3, 4].

Determining the incremental cost effectiveness of the strategies to deliver precision medicine (hereafter short- ened to precision medicine) is important because diverting funding to such newer interventions will involve the real- location of resources from other areas of medicine. The reallocation of funding may affect the health outcomes for relevant populations of patients, representing the opportunity cost of the new intervention. Economic evaluations provide a structured framework to provide evidence sup- porting whether the introduction of precision medicine is an effective use of healthcare resources.

There is some economic evidence supporting exem- plars of precision medicine [5]. However, even with such economic evidence, the introduction of precision medicine into health systems has been slower than anticipated, potentially due to the volume of eligible patients. There is emerging anecdotal, and some empirical, evidence of factors limiting the uptake of precision medicine. In 2015, the charity Cancer Research UK (CRUK) published a report highlighting the significant delays in providing genetic mutation testing to patients [6] and estimated that in the previous year approximately 3500 patients may have missed out on receiving a medicine that may have improved their quality and length of life because of the absence of relevant mutation testing. Questions about the ability of the UK National Health Service (NHS) to implement one of the case study examples of precision medicine, EGFR mutation testing and gefitinib, a spe- cific epidermal growth factor receptor (EGFR) inhibitor licenced for patients with EGFR mutation-positive non- small cell lung cancer, without “substantial investment in time and resources” have been raised. It was believed that the NHS did have capacity to introduce such testing in the short timeframe [7]. However, this proved not to be the case, with the CRUK report suggesting that only 52% of patients received EGFR mutation testing even 4 years after the approval of gefitinib by the National Institute for health and Care Excellence (NICE) [6]. This case study introduces the potential for capacity constraints to be a key barrier to the introduction of precision medicine, even when these have been shown to be cost effective in clinical practice.

No consensus definition of what constitutes a capacity constraint in the context of healthcare interventions exists in the literature. The Oxford English Dictionary (OED) defines capacity as “the amount that something can produce” [8]. This definition implies that it is necessary to be clear what is being produced. The capacity of a healthcare system could therefore be defined as its ability to produce some defined output. In keeping with the extra-welfarist normative under- pinning, assumed by decision-making bodies such as NICE, the relevant output of a healthcare system has been defined as ‘health status’ measured using the quality-adjusted life-year (QALY). The definition of a constraint is “something that controls what you do by keeping you within partici- lar limits” [9]. Combining the definitions of capacity and constraint can be used to propose a working definition of a capacity constraint in a healthcare system which has the goal of maximising health status: ‘Any factor which impedes or limits the amount of health status produced for a popula- tion of patients receiving specified interventions, or policies, provided by the healthcare system’.

The introduction of any healthcare intervention may be impeded by capacity constraints in a healthcare system, which may be particularly extensive and significant for examples of precision medicine due to their nature as complex interventions involving both a test and treatment element [10]. There is some qualitative evidence describing the type of capacity constraints directly relevant to the uptake of precision medicine into practice, including a lack of laborato- ries providing tests, poor logistics resulting in slow test turnaround, a lack of training for clinicians, and insufficient funding for testing or treatments [11–13]. The impact of such capacity constraints on the incremental cost effectiveness of precision medicine has not been well-described. This study aimed to identify a sample of published economic evalua- tions of precision medicine and describe if, and how, these economic evaluations had qualitatively discussed and quantitatively accounted for capacity constraints in the analysis.

Methods

This study used a two-stage systematic review conducted and reported in accordance with published guidelines and reporting criteria [14]. A published search of PubMed by
Payne and colleagues [5] suggested that there were a substantive number of previously published systematic reviews of economic evaluations of precision medicine and related areas such as personalised medicine, pharmacogenetics and pharmacogenomics. Therefore, a de novo systematic review that aimed to identify all previous economic evaluations of precision medicine would overlap significantly with this previous body of work, requiring significant resources in terms of researcher time but unlikely to yield substantively different findings. This study therefore used a strategy to identify and collate published systematic reviews of economic evaluations of precision medicine. From these systematic reviews a sample of economic evaluations of test and treatment-based examples of precision medicine were identified. The identified sample of economic evaluations were published between 2007, when the use of terms related to precision medicine began to occur regularly in the literature [2], and February 2017.

For the purpose of this review precision medicine was defined as an intervention that uses, for example, a test to “identify subgroups of patients with distinct mechanisms of disease, or particular responses to treatments” [1]. Related areas included were therefore as follows: precision medicine; stratified medicine; individualised medicine; genetic medicine; genomic medicine; personalised medicine; and targeted medicine [15]. This systematic collation of systematic reviews published up to February 2017 was then used to identify an exemplar sample of individual economic evaluations and this sample was used to identify if, and how, capacity constraints had been included in the published analyses.

The review involved two stages. Stage one involved the systematic collation of a sample of systematic reviews of economic evaluations of precision medicine. Stage two involved creating a list of relevant individual economic evaluations of precision medicine.

**Stage One**

Table 1 summarises the inclusion criteria used to guide the relevance of published systematic reviews of economic evaluations of precision medicine. To be classified as a systematic review, the published study must have used a systematic approach to search databases of published literature with the aim of identifying all studies that addressed a specified research question.

### Search Strategy

The MEDLINE (inception year: 1946) and EMBASE (inception year: 1980) databases were searched from database inception to February 2017, using an electronic search strategy, to identify all systematic reviews of economic evaluations of precision medicine as defined by this review. The electronic search strategy for this systematic review was based on a published economic evaluation search filter developed by the University Of York-based Centre for Reviews and Dissemination (CRD) [16] and combined with terms relevant to precision medicine and a systematic review search filter, which were informed by a published strategy [5] (see Electronic Supplementary Material (ESM) Appendix 1). In addition, a selection of lead or senior (last) authors of published systematic reviews of economic evaluations of precision medicine were contacted by e-mail to determine whether they knew of other published systematic reviews not identified in the initial search relevant to the selected time period.

### Selection Process

The abstracts identified in the electronic literature search were screened for relevance and inclusion in the review by a team of three reviewers at the Manchester Centre for Health Economics (SJW, Hunter Moore and Sean Gavan). Each abstract was screened by two of the three reviewers. Disagreements on whether a study should be included were resolved by a fourth researcher (Niall Davison).

### Data Extraction and Analysis

The number of systematic reviews of economics evaluations of precision medicine was recorded and their key details summarised using a table (ESM Appendix 2). The identified

<table>
<thead>
<tr>
<th>Table 1 Inclusion criteria for systematic reviews of economic evaluations of precision medicine</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aspect of study</td>
</tr>
<tr>
<td>Population</td>
</tr>
<tr>
<td>Intervention</td>
</tr>
<tr>
<td>Comparator</td>
</tr>
<tr>
<td>Outcomes</td>
</tr>
<tr>
<td>Study type</td>
</tr>
<tr>
<td>Availability</td>
</tr>
</tbody>
</table>

268
systematic reviews were listed and then categorised into one of four categories depending on their stated focus of the systemic review of economic evaluations: (1) test and treat interventions across disease areas; (2) test only interventions across disease areas; (3) test and treat interventions within a given disease area; and (4) test only interventions within a given disease area. Details regarding the number of primary economic evaluations cited in each review were also recorded. These findings were then presented using a narrative summary.

Stage Two

Table 2 summarises the inclusion and exclusion criteria for the individual studies to define a sample of economic evaluations of test and treatment-based precision medicine. Studies that focused on interventions which provide diagnostic information that has no impact on treatment were excluded from this study. A second restriction applied because of the volume of individual economic evaluations limited the identified studies to those published in the 10 years prior to the date of the search.

Search Strategy

A list of individual economic evaluations of precision medicine were identified from the reference lists of the systematic reviews identified in stage one. Grey literature studies that had been identified in the systematic reviews were retained if they met the inclusion criteria for individual studies. This involved a manual search facilitated using a database created in Microsoft Excel® 2010[17].

Selection Process

To determine the relevance of the individual studies and inclusion of primary economic evaluations of test and treatment strategies, the abstracts of the identified studies were imported into the bespoke database in Excel® and then double screened by two reviewers (SJW and Martin Eden, Manchester Centre for Health Economics), with disagreements resolved by a third reviewer (KP).

To identify studies that had discussed capacity issues, a manual keyword search of the PDF for each included study was conducted. To identify relevant keywords, a search of the title and abstracts of relevant theoretical papers was conducted using the Mendeleey reference management software [18–33]. The identified terms were capacity, barrier*, constrain*, restrict*, short (for short-run or short-term), implement*, learn*, inefficien*, bottleneck, scale, utilis*, utiliz*.

Data Extraction and Analysis

The individual studies were initially collated into one of four categories according to the aim of the parent systematic review (as outlined in the stage one data extraction). It was clear, however, that some systematic reviews had identified some studies as test–treat when they were test-only strategies (and vice versa). Therefore, after review of the titles and abstracts of the individual studies, they were then reclassified into one of three categories: (1) economic evaluations of test and treatment interventions; (2) economic evaluations of test-only interventions; and (3) other studies that did not meet the eligibility criteria for this systematic review. Studies classified into the second and third categories were then excluded from this review.

The total number of papers identified and their characteristics are summarised in ESM Appendix 3. Data were extracted by one reviewer (SJW) using a data extraction table produced in Microsoft Word® [34]. Data extraction fields included author; year; country; intervention and comparator; whether the study mentioned capacity, and a brief extract where this was so; and whether the study attempted to account for capacity constraints, and a brief description of the method to account for capacity constraints. Due to the large size of this study, the data extraction table provided only reported papers which as a minimum criteria discussed capacity issues in a qualitative manner (ESM Appendix 4). Studies that attempt to quantitatively account for capacity constraints were summarised using the

<table>
<thead>
<tr>
<th>Table 2 Inclusion criteria for primary economic evaluations of precision medicine</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aspect of study</td>
</tr>
<tr>
<td>Population</td>
</tr>
<tr>
<td>Intervention</td>
</tr>
<tr>
<td>Comparator</td>
</tr>
<tr>
<td>Outcomes</td>
</tr>
<tr>
<td>Study type</td>
</tr>
<tr>
<td>Availability</td>
</tr>
<tr>
<td>Timeframe</td>
</tr>
</tbody>
</table>
Consolidated Health Economic Evaluation Reporting Standards (CHEERS) statement [35]. Themes in the discussion and analysis of capacity issues in the individual evaluations were identified using thematic analysis of the published text in the manuscript. These themes were then discussed in a narrative summary. The narrative summary described the capacity constraints identified in the literature, the extent of the problem of capacity in economic evaluations, and the methods used to deal with these issues in economic evaluations of precision medicine.

**Results**

A total of 45 systematic reviews of economic evaluations of precision medicine published up to and including February 2017 were included in this review (ESM Appendix 2). A Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) diagram showing the process of identifying the systematic reviews for inclusion in this study is presented in Fig. 1. The initial literature search identified 3304 potentially relevant papers in the MEDLINE database and 990 in the EMBASE database. Microsoft Excel® was used to remove duplicate copies first by abstract (n = 283) and secondly by title (n = 92), leaving 3919 papers to review. During double screening of the abstracts of the identified papers, 3871 papers were removed: 3242 were not systematic reviews, 329 did not focus on precision medicine, 133 were not in English, 105 were duplicates, and 62 were not economic evaluations. This left 48 systematic reviews of economic evaluations of precision medicine, to which an additional single study was added following direct contact with key authors in the area. Three reviews were subsequently excluded from data extraction because they did not report the citations for the individual studies included in their review, and one study was removed as on closer inspection it became clear it was not a systematic review of economic evaluations.

A total of 222 individual economic evaluations of precision medicine involving a test–treat strategy published between 2007 and 2015 were included in this review (see Fig. 2). Extraction of the individual economic evaluations from the identified systematic reviews yielded a list of a total of 1101 studies. From this group 477 papers were duplicates and a further 258 papers were removed as they did not evaluate an intervention relevant to precision medicine as defined in this review. Of the remaining 366 studies, a further 62 studies that reported an economic evaluation of a test-only strategy that did not inform a treatment option were excluded. Restricting the individual economic evaluations of test–treat strategies to those published within the last 10 years (2007–2017) yielded 259 studies. Of these studies

![Fig. 1 Identification of systematic reviews of economic evaluations of precision medicine](image)
a further 37 studies were removed during data extraction as it was clear on reading the full text that the study did not focus on precision medicine. Such studies were economic evaluations of medicines which had, in theory, a precision application but were not being evaluated in this way because the costs and consequences of the ‘test’ element were not included in the evaluation. For example, some studies evaluated the cost effectiveness of erlotinib for all non-small cell lung cancer patients rather than just those with EGFR mutations [36, 37].

Description of Systematic Reviews

A total of 45 systematic reviews of economic evaluations of precision medicine were identified which were published between 2004 and February 2017. There were four categories of reviews of economic evaluations: test and treat interventions across disease areas or technology (n = 13); test and treat interventions in a specific disease area of technology (n = 17); diagnostic-only interventions across disease areas or technologies (n = 9); and diagnostic-only interventions in a specific disease or technology area (n = 6). The size of the reviews ranged from zero included papers, for a review of evaluation of genetic diagnosis of aneuploidy in all chromosomes [38], to 140 papers for a review of economic evaluations of targeted and non-targeted therapies for breast cancer [39].

Description of Individual Economic Evaluations

Of the 222 identified economic evaluations of precision medicine, 159 (71%) used QALYs) as the outcome of interest and 45 (20%) used clinical outcomes. A further 19 (9%) studies used a mixture of outcomes. The majority of the economic evaluations were decision-analytic model based (n = 203, 91%), and 20 (9%) were based on primary data from trials. Of the studies using model-based evaluations, 124 (56%) used Markov models, 33 (15%) used decision
trees, 17 (8%) used linked decision trees and Markov models, and 16 (7%) were individual-level simulation models. Six studies used other methods ranging from simple quantitative calculations based on literature reviews and meta-analysis to more complex methods such as dynamic life-cycle modelling, a system dynamics-based approach [40]. In ten (4%) studies it was unclear what modelling approach was used.

The individual economic evaluations identified the costs and consequences of precision medicine for 32 distinct conditions or groups of conditions (Table 3). In total, precision medicine targeting cancer comprised 67% (n = 151) of the individual studies. Evaluations of precision medicine targeting breast cancer dominated the identified economic evaluations (n = 104, 46%). The other common conditions of focus were cardiovascular conditions (n = 22, 10%), including acute myocardial infarction, atrial fibrillation and acute coronary syndrome. Collectively, the interventions in these studies were commonly aimed at preventing strokes. Other more commonly considered conditions included lung cancer (n=18, 8%), colorectal cancer (n=15, 7%) and HIV (n=12, 5%).

**Inclusion of Capacity Constraints**

Of the 222 individual economic evaluations included in this review, 33 mentioned the potential impact of capacity constraints on the costs and consequences of precision medicine in a qualitative sense and nine of these 33 studies went further and attempted to quantify the impact of capacity constraints in the analysis (Table 3). ESM Appendix 4 summarises the 33 included studies. These 33 studies raised key elements related to capacity constraints and how these may impact on the cost effectiveness of precision medicine. These capacity constraints were grouped into four themes: budget constraints; quality of the testing process; ease of test use in clinical practice; and the need for economic evidence to reduce decision uncertainty.

**Budget Constraints**

A key concern relevant to capacity constraints was the total impact of precision medicine on healthcare budgets (budget impact). Budget impact was mentioned in terms of the specific treatment or associated testing and how this may potentially inhibit the uptake and use of these interventions in clinical practice. Three studies focused on the budget impact in broad terms [41–43], but a further ten studies specified the discussion about the financial impact of testing or treat-ment [42, 44–52]. Kondo et al. [47] stressed the need for concern regarding the financial impact of implementing 12,000 new 21-gene assays for breast cancer per year. The considerably larger eligible patient populations who could benefit from interventions such as warfarin dosing tests were suggested to pose more substantial problems in terms of financial impact. Patrick et al. [53] identified that up to 10 million people could benefit from tests costing between US$400 and US$550 per test.

When evaluating the implementation of the more advanced fluorescent in situ hybridisation (FISH) testing for reflex testing human EGFR 2 (HER2) amplifications in breast cancer, Garrison et al. [45] stated that it was unclear whether payers would finance the more rigorous test over the commonly used immunohistochemistry test (IHC). Retél et al. [54] assessed that there was a 75% chance that a 70-gene signature test for breast cancer would not be immediately reimbursed by payers, thereby limiting the availability of testing to patients.

The introduction of genetic testing for breast cancer was seen by the Medical Advisory Secretariat of the Ministry of Health and Long-Term Care, in Ontario, Canada, as an action that could either release health system resources by reducing the number of women receiving unnecessary chemotherapy, or place an additional burden through additional test costs [51]. This report highlighted potential economic incentives that may impact the level of usage of testing, including pharmaceutical companies’ desire to maintain profits in providing the drug to wider population and pay-ers’ desire to reduce drug expenditure.

The cost of specific treatments was also commonly highlighted as a potential barrier to their full implementation. Ito et al. [46] suggested that the out-of-pocket costs of accessing aromatase inhibitors for breast cancer may cause under-utilisation of the potentially health-improving treatment in the USA. It was determined that improving coverage of these drugs by Medicare, a government health insurance scheme, would improve health outcomes while reducing healthcare resource use. Lidgren et al. [49] suggested that individual clinicians may decide not to prescribe treatments due their high cost. The authors also suggested that if different healthcare clinics have different financial budgets then provision of trastuzumab for breast cancer may be variable and this could lead to inequitable access to treatment.

**Quality of Tests and Testing Processes**

A commonly mentioned issue linked to capacity constraints and the cost-effective use of precision medicine was the quality of the tests and testing process to determine the appropriateness of treatment. Suboptimal testing may result from factors including a limited supply of trained test providers, such as pathologists and geneticists, and testing facilities, or unclear reporting of test results to prescribing clinicians. The reduced test quality and potential volume of
### Table 3 Summary of included studies using the Consolidated Health Economic Evaluation Reporting Standards (CHEERs) checklist

<table>
<thead>
<tr>
<th>Study (year)/country</th>
<th>Intervention and comparator</th>
<th>Study population</th>
<th>Economic evaluation type</th>
<th>Evaluation vehicle (model type if applicable)</th>
<th>Time horizon (discount rate)</th>
<th>Analysis</th>
<th>Approach to quantifying capacity constraints</th>
</tr>
</thead>
<tbody>
<tr>
<td>Delea et al. (2012)</td>
<td>Intervention: lapatinib and capcitabine Comparator: capecitabine monotherapy</td>
<td>Women with HER2-positive metastatic breast cancer who have previously received trastuzumab</td>
<td>Cost–utility analysis</td>
<td>Model (partitioned survival analysis)</td>
<td>5 years (3.5%) Incremental analysis reported: yes PSA: yes Other sensitivity analysis: one-way deterministic</td>
<td>Incremental analysis</td>
<td>An average drug wastage number was used in the analysis and this was set to zero in sensitivity analysis, reducing total costs. This suggests cost effectiveness of intervention depends on implementation</td>
</tr>
<tr>
<td>Delea et al. (2013)</td>
<td>Intervention: lapatinib and letrozole Comparator: trastuzumab and anastrozole or trastuzumab alone or letrozole alone</td>
<td>Women with hormone receptor- and HER2-positive metastatic breast cancer</td>
<td>Cost–utility analysis</td>
<td>Model (partitioned survival analysis)</td>
<td>10 years (3.5%) Incremental analysis reported: yes PSA: yes Other sensitivity analysis: one-way deterministic</td>
<td>Incremental analysis</td>
<td>An average drug wastage number was used in the analysis and this was set to zero in sensitivity analysis, reducing total costs. This suggests cost effectiveness of intervention depends on implementation</td>
</tr>
<tr>
<td>Djalalov et al. (2014)</td>
<td>Intervention: EML4-ALK fusion testing and first-line crizotinib treatment Comparator: cisplatin and gemcitabine</td>
<td>Patients with advanced ALK-positive NSCLC</td>
<td>Cost–utility analysis linked to Markov model</td>
<td>Model (decision tree linked to Markov model)</td>
<td>Lifetime (5%) Incremental analysis reported: yes PSA: no Other sensitivity analysis: One-way and two-way deterministic</td>
<td>Incremental analysis</td>
<td>Decision tree includes a branch for whether there is an adequate tissue sample and if not allows for a second biopsy to be taken. It is not clear if these probabilities were varied in sensitivity analysis but the cost of re-biopsy was allowed to vary</td>
</tr>
<tr>
<td>Study (year)/country comparator</td>
<td>Intervention type</td>
<td>Study population</td>
<td>Economic evaluation type</td>
<td>Time horizon (discount rate)</td>
<td>Analysis Approach to quantify capacity constraints</td>
<td>Approach to quantify capacity constraints</td>
<td></td>
</tr>
<tr>
<td>--------------------------------</td>
<td>------------------</td>
<td>------------------</td>
<td>--------------------------</td>
<td>----------------------------</td>
<td>-----------------------------------------------</td>
<td>---------------------------------------------</td>
<td></td>
</tr>
<tr>
<td>Lorenzana et al. (2012) [50]/South Africa</td>
<td>Intervention: genotype assay for selection of third-line ART Comparator: all patients receive second-line treatment or all patients receive third-line treatment</td>
<td>ART-naïve cohort of patients with HIV</td>
<td>Cost-effectiveness analysis</td>
<td>Time horizon not stated (3%)</td>
<td>Incremental analysis reported: yes PSA: no Other sensitivity analysis: one-way and multi-way deterministic</td>
<td>Test cost was varied in sensitivity analysis with suggestions that higher test cost could represent cost when investment is accounted for. No impact on cost effectiveness found</td>
<td></td>
</tr>
<tr>
<td>McCowan et al. (2013) [73]/UK</td>
<td>Intervention: high adherence (≥ 80%) to tamoxifen Comparator: low adherence (&lt; 80%) to tamoxifen</td>
<td>Women with breast cancer</td>
<td>Cost utility analysis reported: yes PSA: yes Other sensitivity analysis: one-way deterministic</td>
<td>Lifetime (3.5%)</td>
<td>Incremental analysis</td>
<td>Evaluation conducted across subgroups of patients with under or over 80% adherence. Low adherence associated with expected loss of 1.12 discounted QALYs and increase of £5970 in medical costs. Methods could be extrapolated to compare cost effectiveness of high patient access to treatments</td>
<td></td>
</tr>
<tr>
<td>Retèl et al. (2012) [54]/The Netherlands</td>
<td>Intervention: 70-gene MammaPrint assay to guide adjuvant breast cancer treatment Comparator: adjuvant! Online algorithm to guide treatment</td>
<td>Women with breast cancer</td>
<td>Cost utility analysis (decision tree and Markov with multiple cohorts and varying parameters)</td>
<td>15 years [4% (costs) and 1.5% (outcomes)]</td>
<td>Incremental analysis reported: yes PSA: no Other sensitivity analysis: none</td>
<td>The researchers modelled the cost effectiveness over time and diffusion of the technology. They include a range of potential scenarios and barriers which affect the diffusion of the technology</td>
<td></td>
</tr>
</tbody>
</table>
To effectively introduce a new test into clinical practice, the relevant health professionals must be aware of the test and have sufficient health training to offer it and guide clinical practice. Insufficient human capital can therefore act as a capacity constraint to moving precision medicine into practice. When tests and treatments are not available tests may both serve to reduce the potential health benefits of precision medicine and act as a barrier to making an informed decision.

### Table 3 (continued)

<table>
<thead>
<tr>
<th>Study population</th>
<th>Economic evaluation type</th>
<th>Evaluation vehicle (model type if applicable)</th>
<th>Time horizon (discount rate)</th>
<th>Analysis</th>
<th>Approach to quantifying capacity constraints</th>
</tr>
</thead>
<tbody>
<tr>
<td>Patients with NSCLC</td>
<td>Cost utility</td>
<td>Model (discrete event simulation)</td>
<td>2 years (3%)</td>
<td>Incremental analysis reported: yes</td>
<td>Includes a parameter for turnaround time and inadequate tissue sample leading to re-biopsy as well as proportion of patients with inadequate tissue sample.</td>
</tr>
<tr>
<td>Romanus et al. [57] USA</td>
<td>Intervention: multiplexed testing for EGF and ALK mutations to guide NSCLC treatment</td>
<td>Comparator: no testing and treatment with pemetrexed and cisplatin</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Vanderlaan et al. [52] USA</td>
<td>Women with node-positive, early-stage breast cancer</td>
<td>Intervention: 21-gene assay to guide adjuvant chemotherapy</td>
<td></td>
<td>Incremental analysis reported: yes</td>
<td></td>
</tr>
<tr>
<td>Comparator: treatment</td>
<td>Cost utility</td>
<td>Model (decision tree)</td>
<td>30 years (3%)</td>
<td>PSA: no</td>
<td>Sensitivity analysis included variations in utilisation rates of testing, although marginal costs were linear so no impact on analysis.</td>
</tr>
</tbody>
</table>

### Ease of Test Use in Clinical Practice

To effectively introduce a new test into clinical practice, the relevant health professionals must be aware of the test and have sufficient health training to offer it and guide clinical practice. Insufficient human capital can therefore act as a capacity constraint to moving precision medicine into practice. When tests and treatments are not available they may both serve to reduce the potential health benefits of precision medicine and act as a barrier to making an informed decision.
assessed in trials and economic evaluation the assumption is that they are used in an optimal, or near optimal, way, and with full adherence to their recommended use by clinicians. However, clinical practice is complex and the way in which a test fits into the current care pathway is not always clear. Nine studies highlighted issues with the transition of tests from research to clinical settings [41, 42, 54, 55, 60–64]. For example, Breijer et al. [61] developed and evaluated two multivariable models to produce an algorithm to predict the risk of endometrial cancer based on patients’ characteristics. While this analysis suggested a reduction in the cost of diagnosing the cancer, these authors proposed that due to the minimal expected value per patient, the algorithm would have to be made very easy to use by clinicians for it to be implemented in clinical practice. The potential reluctance of clinicians to change their prescribing behaviour was shown in a UK study of thio-purine-methyl transferase (TMPT) genotyping to inform the use of azathioprine and the appropriate dosing strategy [62]. Even though the test could be used to better predict which patients would experience profound neutropenia as a reaction from azathioprine, clinicians did not use the test to change the dosing of azathioprine. The authors hypothesised that this is because the clinicians chose to remain conservative with azathioprine prescription due to the other potential adverse reactions that the medicine could cause. Retèl et al. [54] predicted that the use of a 70-gene signature in breast cancer patients would be delayed by “hesitant adopters” who would not use the results of the assay in decision-making. However, these authors included a scenario in their analysis that assumed an increase in the ease of use of the 70-gene signature would subsequently increase the use of test results in clinical decision-making, resulting in improved cost effectiveness of the intervention [54].

Two studies suggested that the willingness of clinicians to use new tests and their ability to effectively use results in decision-making may evolve over time as a result of a learning process [63, 65]. Klang et al. [63] found that clinicians using the Oncotype DX panel did not register treatment decisions for the first 55 of 368 patients as the clinicians were “learning about the technology and how to interpret the results”. The authors of a US study investigating the use of a multi-gene assay in breast cancer explicitly highlighted that previous analyses of Oncotype DX were based on the assumption that the test would be used as dictated by guidelines [65]. They further stated they believe policy makers were interested in “learning how the assay affected outcomes and costs compared with actual practice and after some period of experience with the assay”. Both of these studies indicated that a lower utilisation of testing by health professionals may impact on the incremental cost effectiveness of precision medicine in clinical practice.

Decision Uncertainty

When a new precision medicine is introduced, decision makers, such as individual clinicians, local hospital trusts or national health technology assessment (HTA) agencies, must decide under conditions of uncertainty whether to provide the intervention to patients (implementation). There is a significant cost associated with making the wrong decision related to implementation, which could include reduced patient outcomes at an individual level or societal loss of health from funding cost-ineffective interventions. While commonly used methods such as deterministic and probabilistic sensitivity analysis allow for uncertainty in model parameters to be visualised, a wide range of other uncertainties such as methodological and structural uncertainty may be present in the evaluation of precision medicine [66, 67]. Decision uncertainty around implementation can therefore act as a significant capacity constraint to the introduction of new examples of precision medicine.

As a rationale for conducting economic evaluations, many studies highlighted that evidence of the cost effectiveness of an intervention was a requirement for implementation in a clinical setting [68, 69]. As such, a lack of such evidence on incremental cost effectiveness could inhibit the use of precision medicine in clinical practice. In Canada, the Ontario Medical Advisory Secretariat suggested that evidence on the net resource implications of HER2 testing for breast cancer was required before the intervention could be adopted into “dynamic health systems” [51].

The published evidence also recognised that economic evaluations are conducted at a certain stage in the development of the intervention to deliver precision medicine and that the use and cost effectiveness may change in subsequent clinical practice. Retèl et al. [54, 70] highlighted that their results were based on the assumption of full implementation and that an evaluation using real-life scenarios was in progress at the time of publishing this study.

In theory, the availability of economic evidence can help to fuel increased implementation of precision medicine but, in addition, increased implementation can itself provide economic benefits. Rubinstein et al. [71] identified that if decision makers relied on a passive diffusion of BRCA 1/2 genetic testing into the healthcare system, then potentially the economies of scale which could have been generated from a more managed implementation strategy could be forgone. These authors also suggested that increased implementation of an alternative intervention magnetic resonance imaging in breast cancer may create additional costs and benefits if the imaging becomes a part of the cancer management pathway alongside BRCA 1/2 genetic testing. This
synergistic effect is known as an economy of scope, whereby the same resource can be used to produce or improve multiple services to provide better outcomes.

**Quantifying the Impact of Capacity Constraints in Economic Evaluations**

Nine (see Table 3) of the identified 222 economic evaluations of precision medicine included in this systematic review used, or suggested, techniques that explicitly quantified the impact of capacity constraints [44, 50, 54, 57, 72-76]. All of these studies used decision-analytic models that allowed for sub-perfect implementation of technologies, limiting their potential benefit to society. In other words, for various reasons, fewer than 100% of the eligible patient population were assumed to receive the intervention delivering precision medicine or the intervention was not given in the optimal way, resulting in higher costs or lower benefits than were potentially achievable. In some cases this less than perfect implementation was due to capacity issues such as budget constraints [54], regulation barriers [54] and long test turnaround times [57]. In two other instances, the imperfect implementation was associated with low uptake of the test or treatment [72, 73]. While adherence to medications may appear to be a demand-side problem, it is in fact a complex issue and “the attributes of the health-care system and service delivery may also influence adherence” [73]. For example, if clinicians have limited time to spend with patients, the opportunity to provide effective information about the benefits of adherence and approaches to coping with adverse effects may be limited. The methods used to account for capacity constraints can be categorised as static or dynamic and are described in Sects. 3.4.1 and 3.4.2.

**Static Methods to Account for Capacity Constraints**

Static methods refer to methods used to produce a single cost-effectiveness estimate which takes account of imperfect implementation for one cohort of patients. This estimate of the incremental cost effectiveness may differ from the estimated incremental costs and consequences for a perfectly implemented precision medicine. For example, Delea et al. [74, 75] accounted for drug wastage in their model-based economic evaluation of lapatinib and letrozole for women with HER2-positive breast cancer. This analysis identified that reducing wastage of trastuzumab from 15 to 0% resulted in lapatinib and letrozole being cost effective at £22,895 per QALY gained when compared with trastuzumab, which had been a cost-saving intervention in the base-case analysis. Romanus et al. [57] accounted for specific capacity constraints to multiplexed biomarker testing for non-small cell lung cancer in a sensitivity analysis. When turnaround time increased by a factor of 1.5, the most cost-effective approach changed from “test and treat” to “empiric therapy”, in which patients began treatment with a general chemotherapy agent while waiting for test results. Reducing the proportion of patients being tested from 100% to 5.7% did not impact the rank ordering of the incremental cost effectiveness of the interventions. Likewise, in a study by Vanderlaan and colleagues [72] assumptions about the differences in the uptake of a 21-gene assay for breast cancer did not impact on the incremental cost effectiveness of the intervention. In two further studies, assumptions about the cost of the test or biopsy did not affect the incremental cost effectiveness of a genotype assay for drug selection in patients with HIV [50] or EML4 (echinoderm microtubule associated protein-like 4)-ALK (anaplastic lymphoma kinase) testing for patients with non-small cell lung cancer [44].

McCowan et al. [73] found that adherence to tamoxifen in women with breast cancer significantly impacted the incremental cost effectiveness. On average, patients with an adherence of less than 80% to the treatment were expected to experience 1.12 fewer QALYs than those patients assumed to have over 80% adherence. Furthermore, such patients were expected to experience significantly higher medical costs (£5970, 95% confidence interval 4644–7372).

**Dynamic Methods to Account for Capacity Constraints**

Dynamic methods refer to methods that account for capacity constraints that allow the impact of barriers or constraints to change over time and/or in multiple patient cohorts. As a result, the cost effectiveness of the technology also potentially changes over time. The most comprehensive investigation of capacity constraints and impact on implementation in an economic evaluation using dynamic methods was conducted by Retêl and colleagues [54] who investigated the future potential uptake of a 70-gene signature test in breast cancer. This study used an analytical approach called scenario drafting to identify potential barriers and facilita-tors to the implementation of the technology. These authors constructed different sets of model parameters to reflect the implications of these scenarios, accounting for their perceived likelihood of occurring in reality. These scenarios comprised potential capacity issues such as a lack of reimbursement, testing, regulation issues, uncertainty in the clinical utility of the test, and a lack of use by clinicians due to the difficulty interpreting the tests. The authors also investigated the potential impact of using the test at different stages of the clinical pathway.

The incremental cost effectiveness of the 70-gene signature test was evaluated at different timepoints and implementation levels for three key scenarios: reducing technical failure rates of the test over time; decreasing non-compliance with discordant test results; and increasing financial
reimbursement and clinicians’ uptake of testing. The cost and consequences achieved by testing an additional patient was allowed to vary. As each scenario involved the intervention diffusing into clinical practice at different rates, the costs, health outcomes and cost effectiveness of the intervention differed by year and scenario. A key finding of the study was that the intervention would only be cost effective if use of the test results by clinicians improved over time. At the initial time point, 2005, the intervention was always cost ineffective with an incremental cost-effective-ness ratio (ICER) of €1.9 million. If uptake did not improve, then the ICER only improved to €1.5 million in the best-case scenario. In the scenario where other factors remained equal and uptake improve from 3% in 2005 to 50% in 2010 and 92% in 2010, the ICER reduced from €1.9 million to €26,145 and €11,123, respectively. Failing to overcome the capacity constraints of low financial reimbursement of the test and a lack of uptake by clinicians would therefore mean a potentially cost-effective intervention should not be provided as it would result in a societal loss of health to the population. Removal of the capacity constraints makes the intervention cost effective, but the cost of such capacity investments would also need to be accounted for in an economic evaluation.

Discussion

Precision medicine is often reported to offer a cost-effective approach for the management of a selection of different diseases by using a stratification mechanism to identify which patients may accrue more benefits in terms of response or avoidance of adverse events. Some economic evidence supporting this premise is provided by the growing number of economic evaluations conducted in the past 10 years. This review identified 45 systematic reviews of economic evaluations of precision medicine, which summarised 367 unique individual economic evaluations. Between 2007 and 2015, some 222 economic evaluations were conducted to identify the costs and consequences of test and treat interventions. The first necessary condition for the adoption of an example of precision medicine is demonstration of whether it is a cost-effective use of healthcare resources using appropriate methods of economic evaluation to identify incremental costs and consequences. However, capacity constraints in a healthcare system have the potential to impact on the estimated incremental cost effectiveness of examples of precision medicine in practice. Only 33 studies of a sample of 222 identified economic evaluations qualitatively discussed the potential impact of capacity constraints for the introduction of examples of precision medicine. The core capacity constraints outlined in these papers included budget constraints; quality of tests and testing processes; ease of test use in clinical practice; and decision uncertainty. Even when interventions appeared cost effective, it was suggested they can pose significant financial burdens on payers who have to rapidly provide access to testing services and potentially expensive treatments.

Despite the potential for capacity constraints to affect the incremental cost effectiveness of precision medicine, only nine (4%) economic evaluations from the sample of 222 sought to quantify the effects of limited capacity. All of the nine studies that quantified the effects of limited capacity were based on decision-analytic models rather than trials, and a wide range of model types including decision trees, Markov models and discrete event simulation were used. The presence of these commonly used models in this review suggests that it would be possible to incorporate capacity constraints in many economic evaluations.

While the way in which capacity constraints are included in models will depend on the model type, the methods used could be categorised as either a static or dynamic approach. The static methods included real-world cost-effectiveness analysis and sensitivity analysis in decision-analytic models. Static approaches to incorporating capacity constraints only give a representation of how the constraints impact the cost effectiveness of the intervention for a single cohort of patients at a single point in time. Dynamic approaches to quantifying capacity constraints allow for the fact that health system capacity can change over time and this can have a changing impact on the cost effectiveness of the intervention. The example of Retel et al. [54] focused on how the level of uptake impacts on the cost effectiveness of the intervention. However, the method used could also be applied to capacity constraints. If the ICER is non-linear and depends on the level of implementation, then any factor that impedes implementation can in theory render the intervention cost ineffective. For example, if numerous repeat EGFR mutation tests are required due to insufficient samples at the start of implementation, this will raise the cost of testing and reduce patient benefits due to delays in receiving treatment. With greater communication regarding biopsy requirements and learning by clinicians, better samples could be obtained, reducing costs and improving benefits. Therefore, the ICER will be dependent on the extent to which the intervention is being effectively implemented and the impact of testing knowledge as a capacity constraint. This study shows the key interaction between barriers that impede the use of an intervention and the potential for marginal costs and benefits to vary depending on the level of implementation. The result is situations where capacity constraints and other barriers to implementation cause the intervention to become cost ineffective. Implementing the precision medicine at this level will result in healthcare resources being diverted away from other areas where they could be put to better use, lowering the overall health of patients in the healthcare system.
The Retèl et al. [54] study also raised an important consideration in that actions can be taken to overcome capacity constraints and to improve the incremental cost effectiveness of precision medicine. For example, testing guidelines could be introduced or education programmes used to improve the quality of biopsy samples for testing. Such an intervention itself would have a cost but would also provide benefits in improving the benefits for patients and making the intervention more cost effective. The evaluation of such strategies to improve the implementation of interventions is known as value of implementation analysis [27]. This systematic review did not identify any value of implementation analyses investigating the implementation of precision medicine. The use of dynamic, multi-cohort decision-analytic models and the value of an implementation approach may provide better evidence than static methods to decision makers in allowing them to understand how best to implement new examples of precision medicine in a cost-effective way by investing in improving health system capacity. In order to forecast the potential capacity constraints before the approval of a precision medicine by an HTA body, the use of qualitative methods such as interviews or focus groups may prove useful. For example, Retèl et al. [54] used the Delphi method with a group of clinicians to identify potential barriers to implementing the MammaPrint test.

**Limitations**

The economic evaluation of precision medicine is an expanding research area, and due to the size of the literature base, some restrictions to the inclusion of papers in this study were required. The use of a search strategy that identified previous systematic reviews reduced the number of papers for abstract screening. Collating these studies provided a comprehensive set of primary studies and theoretically all papers previously published in this area. However, due to the length of time required to conduct a systematic review, it is possible that economic evaluations of precision medicine have been published in the time since the literature searches of the most recent systematic reviews were conducted. This means there may be a gap in the studies identified between 2015 and 2017. This could cause bias in the results of this study if methods to include capacity constraints in economic evaluations have been recently developed and applied. The authors of this review are not aware of any recently published novel methods for accounting for capacity constraints in economic evaluations of precision medicine.

The focus on studies published in the last 10 years and evaluations that focused on test and treatment strategies could also feasibly have excluded studies that discussed capacity. While studies published before 2007 may have discussed capacity due to the novelty of precision medicine, it is unlikely that early economic evaluations of precision treatments would have incorporated complex methodological adjustments for capacity. Capacity could be a significant issue in the provision of precision treatments which only provide diagnostic information. Such studies were excluded from this review.

This review has taken a broad definition of a capacity constraint as any factor which impedes the full benefits of an intervention from being realised. While this includes factors such as finite budgets and the quality of the testing process, it also includes more abstract concepts such as low usage of tests due to a lack of knowledge of the technology amongst clinicians. Some of the included studies also investigated adherence to medicine and uptake of treatment. While low adherence could be due to capacity constraints in a lack of patient education about the benefits of treatment, it may also be due to underlying patient preferences for the treatment or adverse effect profile. Therefore, while full implementation of a precision medicine will rely on a lack of capacity constraints on the supply side, it could also be impeded by low demand for the treatment by patients. This could be the case if a new treatment had a greater risk of more adverse effects, more severe adverse effects or a different range of adverse effects. There may, therefore, be a limit to the level of implementation that can be achieved by investing in capacity.

While this review has focused on the impact of capacity constraints for the economic evaluation of precision medicine, such constraints in the healthcare system may have a significant impact on the cost effectiveness of interventions in other medical areas. For example, Jahn et al. [28] explored the cost effectiveness of drug-eluting stents in the presence of capacity constraints using a discrete event simulation model. Capacity constraints may also be particularly significant for the cost effectiveness of organ transplants where there is a limited availability of donors [77]. In addition, capacity constraints and their impact may be larger in countries with developing healthcare systems [78].

**Conclusions**

The results of this systematic review suggest that a wide variety of capacity constraints could have implications for the cost effectiveness of precision medicine in clinical practice, but the majority of economic evaluations of precision medicine do not account for such constraints. In the studies that did account for limited health system capacity, a variety of methods had been used, with most relying on static comparisons of the cost effectiveness of examples of precision medicine for a single cohort with or without health system capacity constraints.

Studies should account for changing health system capacity, which may have implications for the cost effectiveness of interventions across multiple cohorts in different years.
This is because when combined with varying marginal costs and benefits, implementation-limiting capacity constraints can result in interventions that are cost effective at the population level becoming inefficient in the short-run. Health economists should endeavour to forecast potential barriers to implementing precision medicine and to evaluate potential strategies to invest in capacity. A number of studies have reported qualitative investigations of such barriers [11–13] and combining these approaches with dynamic methods for quantifying such barriers may help to provide decision makers with more robust evidence as to how to cost-effectively implement such interventions and take resource and capacity constraints into account.

Acknowledgements The authors would like to thank Sean Gavan, Hunter Moore and Martin Eden for their help in second screening the abstracts of identified systematic reviews and primary economic evaluations. We would also like to thank Niall Davison for his help in constructing the search strategy for identifying the systematic reviews and resolving disagreements in screening.

Author Contributions SW conceptualised the study, conducted the literature search, screened abstracts, extracted data, analysed the data, wrote the first draft and edited the manuscript. WN conceptualised the study, and reviewed and edited the manuscript. KP conceptualised the study, screened abstracts, and reviewed and edited the manuscript.

Data Availability Statement The list and references of identified systematic reviews of precision medicine are available in Electronic Supplementary Material Appendix 2. The references of included economic evaluations of examples of precision medicine are available in Electronic Supplementary Material Appendix 3.

Compliance with Ethical Standards

Funding This work was supported by a Wellcome Trust Grant (203405) as part of a Society and Ethics Doctoral Studentship for the PhD programme of Stuart Wright. Financial support for this study was provided entirely by this grant from the Wellcome Trust. The funding agreement ensured the authors’ independence in designing the study, interpreting the data, writing and publishing the report.

Conflict of Interest Stuart Wright, William Newman and Katherine Payne declare that they have no conflicts of interest.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

21. Sheldon TA. What’s the evidence that NICE guidance has been implemented? Results from a national evaluation using time series analysis, audit of patients’ notes, and interviews. BMJ. 2004;329:999. https://doi.org/10.1136/bmj.329.7473.999.


## Appendix 2.2: Search strategy

<table>
<thead>
<tr>
<th>MEDLINE</th>
<th>EMBASE</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>CRD MEDLINE Review Terms</strong></td>
<td><strong>CRD EMBASE Review Terms</strong></td>
</tr>
<tr>
<td>1. review.ti,ab.</td>
<td>1. exp meta analysis/</td>
</tr>
<tr>
<td>2. review.pt</td>
<td>2. meta-analysis$.ti,ab.</td>
</tr>
<tr>
<td>3. meta-analysis.ab</td>
<td>3. metaanalysis$.ti,ab.</td>
</tr>
<tr>
<td>4. meta-analysis.pt</td>
<td>4. meta analysis$.ti,ab.</td>
</tr>
<tr>
<td>5. meta-analysis.ti</td>
<td>5. review$.ti.</td>
</tr>
<tr>
<td>6. or/1-5</td>
<td>6. overview$.ti.</td>
</tr>
<tr>
<td>7. letter.pt</td>
<td>7. (synthes$ adj3 (literature$ or research$ or studies or data)).ti,ab.</td>
</tr>
<tr>
<td>8. comment.pt</td>
<td>8. pooled analysis$.ti,ab.</td>
</tr>
<tr>
<td>9. editorial.pt</td>
<td>9. ((data adj2 pool$) and studies).mp.</td>
</tr>
<tr>
<td>10. or/7-9</td>
<td>10. (MEDLINE or medlars or embase or cinahl or scisearch or psychinfo or psychinfo or psychlit or psyclit).ti,ab.</td>
</tr>
<tr>
<td>11. 6 not 10</td>
<td></td>
</tr>
</tbody>
</table>

<p>| EED Economic Terms | |
| 12. economics/ | 11. ((hand or manual or database$ or computer$) adj2 search$).ti,ab. |
| 13. exp &quot;costs and cost analysis&quot;/ | 12. ((electronic or bibliographic$) adj2 (database$ or data base$)).ti,ab. |
| 14. Economics, Dental/ | 13. ((review$ or overview$) adj10 (systematic$ or methodologic$ or quantitativ$ or research$ or literature$ or studies or trial$ or effective$)).ab. |
| 15. exp economics, hospital/ | 14. 1 or 2 or 3 or 4 or 5 or 6 or 7 or 8 or 9 or 10 or 11 or 12 or 13 |
| 19. (economic$ or cost or costs or costly or costing or price or prices or pricing or pharmacoeconomic$).ti,ab. | 18. (patient$ adj2 review$).ti,ab,sh. |
| 21. value for money.ti,ab. | 20. (peer adj2 review$).ti,ab,sh. |
| 22. budget$.ti,ab. | 21. (chart$ adj2 review$).ti,ab,sh. |
| 23. or/12-22 | 22. (case$ adj2 report$).ti,ab,sh. |
| 24. ((energy or oxygen) adj cost).ti,ab. | 23. (rat or rats or mouse or mice or hamster or hamsters or animal or animals or dog or dogs or cat or cats or bovine or sheep).ti,ab,sh. |
| 25. (metabolic adj cost).ti,ab. | 24. 15 or 16 or 17 or 18 or 19 or 20 or 21 or 22 or 23 |
| 26. ((energy or oxygen) adj expenditure).ti,ab. | 25. 14 not 24 |
| 27. or/24-26 | 26. editorial.pt. |
| 28. 23 not 27 | 27. letter.pt. |
| 29. exp animals/ not humans/ | 28. 26 or 27 |
| 30. 28 not 29 | 29. 25 not 28 |
| 31. bmj.jn. | 30. exp animal/ |
| 32. &quot;cochrane database of systematic reviews&quot;.jn. | 31. exp nonhuman/ |
| 33. health technology assessment winchester England.jn | 32. 30 or 31 |
| 34. or/31-33 | 33. exp human/ |
| 35. 30 not 34 | 34. 32 not (32 and 33) |
| Genetic, Genomic and Precision Medicine Terms | 35. 29 not 34 |
| 36. ((genetic$ or pharmacogen$ or pharmaco-gen$) adj2 (test$ or tech$ or assessment or evaluation or intervention$ or screen$ or service$)).mp. | EED Economic Terms |
| 37. biomarker.mp. | |
| 38. (next adj generation adj sequencing).mp. | |</p>
<table>
<thead>
<tr>
<th>40. 60 adj2 sequencing).mp.</th>
<th>36. Health Economics/</th>
</tr>
</thead>
<tbody>
<tr>
<td>41. or/36-40</td>
<td>37. exp Economic Evaluation/</td>
</tr>
<tr>
<td>42. (genom$ or precision or personali$ or stratif$ or individuali$ or target$ or P4) adj (medic$ or treatment or therap$)</td>
<td>38. exp Health Care Cost/</td>
</tr>
<tr>
<td>43. 41 or 42</td>
<td>39. pharmacoeconomics/</td>
</tr>
<tr>
<td>44. 11 and 35 and 43</td>
<td>40. 36 or 37 or 38 or 39</td>
</tr>
<tr>
<td></td>
<td>41. (econom$ or cost or costs or costly or costing or price or prices or pricing or pharmacoeconomic$).ti,ab.</td>
</tr>
<tr>
<td></td>
<td>42. (expenditure$ not energy).ti,ab.</td>
</tr>
<tr>
<td></td>
<td>43. (value adj2 money).ti,ab.</td>
</tr>
<tr>
<td></td>
<td>44. budget$.ti,ab.</td>
</tr>
<tr>
<td></td>
<td>45. 41 or 42 or 43 or 44</td>
</tr>
<tr>
<td></td>
<td>46. 40 or 45</td>
</tr>
<tr>
<td></td>
<td>47. note.pt.</td>
</tr>
<tr>
<td></td>
<td>48. 46 not 47</td>
</tr>
<tr>
<td></td>
<td>49. (metabolic adj cost).ti,ab.</td>
</tr>
<tr>
<td></td>
<td>50. ((energy or oxygen) adj cost).ti,ab.</td>
</tr>
<tr>
<td></td>
<td>51. ((energy or oxygen) adj expenditure).ti,ab.</td>
</tr>
<tr>
<td></td>
<td>52. 49 or 50 or 51</td>
</tr>
<tr>
<td></td>
<td>53. 48 not 52</td>
</tr>
<tr>
<td></td>
<td>54. 0959-8146.is.</td>
</tr>
<tr>
<td></td>
<td>55. (1469-493X or 1366-5278).is.</td>
</tr>
<tr>
<td></td>
<td>56. 1756-1833.en.</td>
</tr>
<tr>
<td></td>
<td>57. or/54-56</td>
</tr>
<tr>
<td></td>
<td>58. 53 not 57</td>
</tr>
<tr>
<td></td>
<td>59. conference abstract.pt.</td>
</tr>
<tr>
<td></td>
<td>60. 58 not 59</td>
</tr>
</tbody>
</table>

Genetic, Genomic and Precision Medicine Terms

| 61. ((genetic$ or pharmacogen$ or pharmacogen$) adj2 (test$ or tech$ or assessment or evaluation or intervention$ or screen$ or service$)).mp. |
| 62. biomarker.mp. |
| 63. (next adj generation adj sequencing).mp. |
| 64. (high adj throughput adj sequencing).mp. |
| 65. (whole adj2 sequencing).mp. |
| 66. or/61-65 |
| 67. (genom$ or precision or personali$ or stratif$ or individuali$ or target$ or P4) adj (medic$ or treatment or therap$) |
| 68. 66 or 67 |
| 69. 35 and 60 and 68 |
# Appendix 2.3: Summary of Included Systematic Reviews of Economic Evaluations of Precision Medicine

<table>
<thead>
<tr>
<th>Author</th>
<th>Year</th>
<th>Intervention</th>
<th>Reported Health Condition</th>
<th>Databases Searched</th>
<th>Studies included (n)</th>
<th>Reference*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Antoñanzas et al.</td>
<td>2012</td>
<td>Genetic tests</td>
<td>Cystic fibrosis; Breast, Colorectal, Ovarian, and Prostate cancer; Hereditary hemochromatosis; Down’s syndrome; Familial hypercholesterolemia; Thrombophilia</td>
<td>PubMed, Euronheed, University of York (HTA), DARE, NHS EED, Scopus</td>
<td>51</td>
<td>[1]</td>
</tr>
<tr>
<td>Assasi et al.</td>
<td>2012</td>
<td>Predictive, diagnostic, or preventive tests, and genetic tests to guide treatment</td>
<td>Hypercholesterolemia; Prenatal chromosome aberration; Lynch syndrome; Breast, Colorectal, and Non-small cell cancers; Hemochromatosis; Fragile X syndrome; Cytomegalovirus</td>
<td>Medline, Embase, Evidence Based Medicines’s HTA database, PubMed, Wiley’s Cochrane Library, HEED, CRDHTA database</td>
<td>15</td>
<td>[2]</td>
</tr>
<tr>
<td>Berm et al.</td>
<td>2016</td>
<td>Pharmacogenetic and pharmacogenomic screening tests</td>
<td>Specific health conditions not reported</td>
<td>PubMed</td>
<td>38</td>
<td>[4]</td>
</tr>
<tr>
<td>Carlson et al.</td>
<td>2005</td>
<td>Genetic services (a disease that was primarily genetic or involved a genetic test)</td>
<td>Breast and Colorectal Cancer; Aneuploidies; Cystic fibrosis; Thrombophilia; Fetal anomalies; Hemochromatosis; Hypercholesterolemia; Hemoglobinopathy; Metabolism errors; Rheumatoid arthritis/Lupus</td>
<td>PubMed, Proquest, LexisNexis, Expanded Academic Index, Harvard Review of Economic Analyses, PsycINFO, NICE, CCOHTA</td>
<td>63</td>
<td>[7]</td>
</tr>
<tr>
<td>D’Andrea et al.</td>
<td>2016</td>
<td>BRCA testing</td>
<td>Breast and ovarian cancer</td>
<td>Medline, Scopus, HEED, EconLit, HTA, NHS EED</td>
<td>9</td>
<td>[9]</td>
</tr>
<tr>
<td>Degeling et al.</td>
<td>2017</td>
<td>Tests or prediction models to stratify patients into subgroups for screening, treatment targeting, or treatment monitoring</td>
<td>Breast, Lung, Colorectal, and Prostate cancer; Cardiovascular disease; HIV; Hepatitis C; Alzheimer’s disease; Atrial fibrillation; Neonatal disease; Rheumatoid arthritis; Depressive disorder, Type 2 diabetes; Acute myeloid leukemia</td>
<td>PubMed</td>
<td>31</td>
<td>[10]</td>
</tr>
<tr>
<td>Authors</td>
<td>Year</td>
<td>Study Title</td>
<td>Conditions</td>
<td>Sources</td>
<td>References</td>
<td></td>
</tr>
<tr>
<td>---------</td>
<td>------</td>
<td>-------------</td>
<td>------------</td>
<td>---------</td>
<td>------------</td>
<td></td>
</tr>
<tr>
<td>Djalalov et al.</td>
<td>2011</td>
<td>Genetic tests or services to detect a heritable disease</td>
<td>Inflammatory bowel disease; Breast and Colorectal cancer; Long QT syndrome; Fragile X syndrome; Schizophrenia; Crohn’s disease; Cystic fibrosis; Hypertension; Nephropathy; Hemochromatosis; Learning disability; Smoking cessation; Aneuploidy; Atrial fibrillation</td>
<td>PubMed, Medline, Proquest, LexisNexis, Expanded Academic Index, The Harvard Review of Economic Analyses, PsycINFO, NICE, CADTH</td>
<td>26 [11]</td>
<td></td>
</tr>
<tr>
<td>Douglas et al.</td>
<td>2016</td>
<td>Population-wide and high-risk genetic screening</td>
<td>Lynch syndrome; Hereditary breast and ovarian cancer; Familial hypercholesterolemia, Romano-Ward long QT syndrome; Brugada syndrome; Hypertrophic cardiomyopathy; Dilated cardiomyopathy; MYH-associated polyposis; Multiple endocrine neoplasia type 2</td>
<td>PubMed</td>
<td>32 [13]</td>
<td></td>
</tr>
<tr>
<td>Ferrusi et al.</td>
<td>2009</td>
<td>Trastuzumab and HER2 testing</td>
<td>Breast cancer</td>
<td>Biosis, Cochrane, HEED, CRD, Embase, EconLit, Medline, PubMed</td>
<td>17 [14]</td>
<td></td>
</tr>
<tr>
<td>Author(s)</td>
<td>Year</td>
<td>Study Details</td>
<td>Disease/Condition</td>
<td>Literature Databases/Database</td>
<td>N</td>
<td></td>
</tr>
<tr>
<td>----------------</td>
<td>------</td>
<td>---------------------------------------------------</td>
<td>-----------------------------------------------------</td>
<td>-----------------------------------------------------------</td>
<td>----</td>
<td></td>
</tr>
<tr>
<td>Ferrusi et al.</td>
<td>2011</td>
<td>Trastuzumab and HER2 testing</td>
<td>Breast cancer</td>
<td>Biosis, Cochrane, CRD, Embase, EconLit, Medline, PubMed, HEED</td>
<td>14</td>
<td></td>
</tr>
<tr>
<td>Frank et al.</td>
<td>2013a</td>
<td>Genomic sequencing technologies</td>
<td>Specific health conditions not reported</td>
<td>Deutsches Ärzteblatt, Biosis Previews, Cochrane Database of Systematic Reviews, DAHTA-Datenbank, Embase, GMS, GMS Meetings, Social SciSearch, Health Technology Assessment Database, SciSearch, Krause &amp; Pachernegg Verlagsdatenbank, Medline, NHS Economic Evaluation Database, Thieme Verlagsdatenbank, and Thieme Verlagsdatenbank PrePrint.</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>Frank et al.</td>
<td>2013b</td>
<td>Certuximab and panitumumab with and without KRAS testing</td>
<td>Colorectal cancer</td>
<td>DIMDI, Medline, Biosis, GMS, DAHTA, Embase</td>
<td>7</td>
<td></td>
</tr>
<tr>
<td>Gavan et al.</td>
<td>2014</td>
<td>Explicit stratified approach to treatment</td>
<td>Rheumatoid Arthritis</td>
<td>Medline, Embase, Web of Science, NHS EED</td>
<td>10</td>
<td></td>
</tr>
<tr>
<td>Giacomini et al.</td>
<td>2003</td>
<td>Molecular genetic tests</td>
<td>Breast and ovarian cancer; Thromboembolic disease; Endocrine neoplasia; Hemochromatosis; Cystic fibrosis; Fragile X syndrome;</td>
<td>Medline</td>
<td>14</td>
<td></td>
</tr>
</tbody>
</table>

288
<table>
<thead>
<tr>
<th>Authors</th>
<th>Year</th>
<th>Description</th>
<th>Condition</th>
<th>Database(s)</th>
<th>Count</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gonzalez et al.</td>
<td>2015</td>
<td>Genetic testing and electrocardiogram testing for diagnosis</td>
<td>Long QT Syndrome</td>
<td>Medline, Embase, CRD</td>
<td>4</td>
<td>[20]</td>
</tr>
<tr>
<td>Hatz et al.</td>
<td>2014</td>
<td>Individualized medicine – “a therapeutic approach tailoring therapy for genetically defined subgroups of patients”</td>
<td>Breast, Colorectal, Lung, and Ovarian cancer; HIV, Hepatitis C; Atrial fibrillation; Thrombosis; Hypercholesterolemia; Hypertrophic cardiomyopathy; Long QT syndrome; Smoking cessation; Acute coronary syndrome; Nephropathies; Depression; Schizophrenia; Periodontal disease; Cystic fibrosis; Acute lymphoblastic leukemia; MYH-associated polyposis; Epilepsy; Pulmonary fibrosis; Kidney failure; Cardiovascular disease; Hypertension</td>
<td>Medline</td>
<td>84</td>
<td>[22]</td>
</tr>
<tr>
<td>Jarrett et al.</td>
<td>2006</td>
<td>Genetic techniques and testing technologies used for the detection or treatment of diseases with known genetic causes or associations</td>
<td>Breast, Ovarian, and Colorectal cancer; Cystic fibrosis; Down’s syndrome; Remaining diseases not reported</td>
<td>PubMed, Embase, Econlit, NHS EED, Office of Health Economics HEED, NHS HTA</td>
<td>37</td>
<td>[23]</td>
</tr>
<tr>
<td>Langer et al.</td>
<td>2010</td>
<td>PET and CT-based strategies with invasive and non-invasive diagnostic strategies</td>
<td>Solitary pulmonary nodule; Malignant melanoma; Nasopharyngeal carcinoma; Ovarian, Colorectal, Head and neck, Breast, and Non-small cell lung cancer</td>
<td>Cochrane Library, DARE, Embase, HTA Database, NHS EED, PubMed, RePEc, Web of Science</td>
<td>14</td>
<td>[26]</td>
</tr>
<tr>
<td>Lansdorp-Vogelaar et al.</td>
<td>2011</td>
<td>Colorectal cancer screening</td>
<td>Colorectal cancer</td>
<td>Medline, Embase, the Cost-Effectiveness Analysis Registry, British National Health Service Economic</td>
<td>32</td>
<td>[27]</td>
</tr>
<tr>
<td>Study</td>
<td>Year</td>
<td>Description</td>
<td>Disease</td>
<td>Resources</td>
<td>Citations</td>
<td></td>
</tr>
<tr>
<td>-------</td>
<td>------</td>
<td>-------------</td>
<td>---------</td>
<td>-----------</td>
<td>-----------</td>
<td></td>
</tr>
<tr>
<td>Lee et al.</td>
<td>2015</td>
<td>Genetic diagnosis for aneuploidy in all 24 chromosomes (PGD-A)</td>
<td>Aneuploidy in all 24 chromosomes (PGD-A)</td>
<td>Medline, Embase, Scopus, Cochrane Library, NHS EED, EconLit</td>
<td>0 [28]</td>
<td></td>
</tr>
<tr>
<td>Lieberthal</td>
<td>2013</td>
<td>Genomic testing</td>
<td>Breast cancer</td>
<td>PubMed, Google Scholar</td>
<td>9 [29]</td>
<td></td>
</tr>
<tr>
<td>Oosterhoff et al.</td>
<td>2016</td>
<td>Biomarkers as a diagnostic (for diagnosis, staging, and as a companion diagnostic)</td>
<td>Colorectal, Lung, Renal, Thyroid, Breast, Endometrial, and Intra-abdominal cancer; Cardiovascular disease; Respiratory disease; Diabetes; Circulatory disease</td>
<td>PubMed, NHS EED</td>
<td>33 [32]</td>
<td></td>
</tr>
<tr>
<td>Parkinson et al.</td>
<td>2014</td>
<td>Trastuzumab</td>
<td>HER2 positive metastatic breast cancer</td>
<td>Medline, Embase, Cochrane Database, EED and HTA</td>
<td>12 [33]</td>
<td></td>
</tr>
<tr>
<td>Payne et al.</td>
<td>2009</td>
<td>TPMT testing for dosing of azathioprine</td>
<td>Crohn’s disease; Rheumatoid arthritis; Systemic lupus erythematosus; Inflammatory</td>
<td>Medline, Embase, PsychInfo, HAPI, CINAHL</td>
<td>6 [34]</td>
<td></td>
</tr>
<tr>
<td>Authors</td>
<td>Year</td>
<td>Study Title</td>
<td>Indications</td>
<td>Database(s)</td>
<td>References</td>
<td></td>
</tr>
<tr>
<td>------------------</td>
<td>------</td>
<td>-------------------------------------------------------------------------------</td>
<td>--------------------------------------------------------------------------------------------------</td>
<td>----------------------------------------------------------------------------</td>
<td>------------</td>
<td></td>
</tr>
<tr>
<td>Plöthner et al.</td>
<td>2016</td>
<td>Targeted therapies using pharmacogenomic and pharmacogenetic testing</td>
<td>Epilepsy; Neuropathic pain; HIV; Inflammatory bowel disease; Rheumatoid arthritis; Systemic lupus erythematosus; Acute lymphoblastic leukemia; Breast, Colorectal, and Non-small cell lung cancer</td>
<td>German Institute for Medical Documentation and Information meta-database</td>
<td>27 [36]</td>
<td></td>
</tr>
<tr>
<td>Plumpton et al.</td>
<td>2016</td>
<td>Pharmacogenetic testing to prevent adverse drug reactions</td>
<td>Acute coronary syndrome; Atrial fibrillation; Acute lymphoblastic leukemia; Crohn’s disease; Cardiovascular disease; Deep vein thrombosis; Inflammatory bowel disease; Major depressive disorder; Myocardial infarction; Colorectal cancer; Peripheral arterial disease; Pulmonary embolus; Rheumatoid arthritis; Stevens-Johnson syndrome; Systemic lupus erythematosus; Ulcerative colitis</td>
<td>Embase, Medline, NHS EED</td>
<td>47 [37]</td>
<td></td>
</tr>
<tr>
<td>Poonawalla et al.</td>
<td>2015</td>
<td>Chemotherapeutic agents and targeted biologics</td>
<td>Ovarian cancer</td>
<td>Medline, PubMed, Embase</td>
<td>28 [38]</td>
<td></td>
</tr>
<tr>
<td>Study</td>
<td>Year</td>
<td>Focus</td>
<td>Conditions</td>
<td>Databases</td>
<td>References</td>
<td></td>
</tr>
<tr>
<td>-------------------------------</td>
<td>------</td>
<td>------------------------</td>
<td>-------------------------------------------------------------------------------------------------</td>
<td>------------------------------------------------</td>
<td>------------</td>
<td></td>
</tr>
<tr>
<td>Rogowski et al.</td>
<td>2006</td>
<td>Genetic screening</td>
<td>Cystic fibrosis; Diabetes; Hereditary Breast and Ovarian cancer; Retinoblastoma; Familial hypercholesterolaemia; Hereditary haemochromatosis; Hereditary nonpolyposis colorectal carcinoma; Familial adenomatous polyposis colorectal cancer</td>
<td>PubMed, Biosys, Cochrane, DAHTA, Embase, IHTA, Medline, NHS-HTA-DARE, NHS-CRD-HTA, NHS-EED, SOMED</td>
<td>[39]</td>
<td></td>
</tr>
<tr>
<td>Smieliauskas et al.</td>
<td>2014</td>
<td>Targeted oral anti-cancer medications</td>
<td>Breast cancer; Chronic myeloid leukemia; Gastrointestinal stromal tumor; Hepatocellular carcinoma; Non-small cell lung cancer; Renal cell cancer; Pancreatic cancer</td>
<td>PubMed, Cochrane Library, NIHR HTA</td>
<td>[40]</td>
<td></td>
</tr>
<tr>
<td>Stevanovic et al.</td>
<td>2012</td>
<td>Risk prediction models</td>
<td>Cardiovascular disease</td>
<td>PubMed, NHS EED</td>
<td>[41]</td>
<td></td>
</tr>
<tr>
<td>Vegter et al.</td>
<td>2008</td>
<td>Pharmacogenetic and pharmacogenomic screening programmes</td>
<td>Specific health conditions not reported</td>
<td>PubMed, Embase, Web of Science</td>
<td>[42]</td>
<td></td>
</tr>
<tr>
<td>Vegter et al.</td>
<td>2010</td>
<td>Pharmacogenetic and pharmacogenomic screening programmes</td>
<td>Thromboembolism; Atrial fibrillation; Dermatologic conditions; Rheumatologic</td>
<td>PubMed</td>
<td>[43]</td>
<td></td>
</tr>
<tr>
<td>Verhoef et al.</td>
<td>2010</td>
<td>Pharmacogenetic-guided dosing of coumarin therapy</td>
<td>Atrial fibrillation, patients starting warfarin</td>
<td>PubMed, Embase, NHS EED, Web of Science</td>
<td>9</td>
<td>[44]</td>
</tr>
<tr>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>Wong et al.</td>
<td>2010</td>
<td>Pharmacogenomics</td>
<td>Breast, Colon, and Lung cancer; Rheumatologic conditions; Gastrointestinal diseases; Thromboembolic conditions; HIV; Nicotine addiction; Nephropathy; Psychiatric conditions</td>
<td>PubMed, NICE, Tufts CEA registry, CADTH</td>
<td>34</td>
<td>[45]</td>
</tr>
</tbody>
</table>
Appendix 2.4: References of all Included Economic Evaluations of Precision Medicine


298


Fleeman N, NIHR Health Technology Assessment Programme (Great Britain), NIHR Evaluation Trials and Studies Coordinating Centre (Great Britain). Lapatinib and trastuzumab in combination with an aromatase inhibitor for the first-line treatment of metastatic hormone receptor-positive breast cancer which over-expresses human epidermal growth factor 2 (HER2) : a systematic review and economic analysis. Heal. Technol. assessment,. 2011;15:xvi, 100 p. 61.


Garrison LP, Veenstra DL. The economic value of innovative treatments over the product life cycle: The case of targeted trastuzumab therapy for breast cancer. Value Heal. [Internet]. International Society for Pharmacoeconomics and Outcomes Research (ISPOR); 2009;12:1118–23.


Holt S, Bertelli G, Humphreys I, Valentine W, Durrani S, Pudney D, et al. A decision impact, decision conflict and economic assessment of routine Oncotype DX testing of 146 women with node-


Millar JA, Millward MJ. Cost Effectiveness of Trastuzumab in the Adjuvant Treatment of Early A Lifetime Model. 2007;25:429–42. 139.


Schackman BR, Scott CA, Walensky RP, Losina E, Freedberg KA, Sax PE. The cost-effectiveness of HLA-B*5701 genetic screening to guide initial antiretroviral therapy for HIV. AIDS. 2008;22. 179.


Skedgel C, Rayson D, Younis T. The cost-utility of sequential adjuvant trastuzumab in women with Her2/Neu-Positive Breast Cancer: An analysis based on updated results from the HERA trial. Value Heal. [Internet]. International Society for Pharmacoeconomics and Outcomes Research (ISPOR); 2009;12:641–8.


Thompson D, Taylor DCA, Montoya EL, Winer EP, Jones SE, Weinstein MC. Cost-effectiveness of switching to exemestane after 2 to 3 years of therapy with tamoxifen in postmenopausal women with early-stage breast cancer. Value Heal. [Internet]. International Society for Pharmacoeconomics and Outcomes Research (ISPOR); 2007;10:367–76.


Appendix 2.5: Location of Economic Evaluations of Precision Medicines (2007-2017)

<table>
<thead>
<tr>
<th>Country</th>
<th>Number of Studies(^1)</th>
<th>Proportion of Studies (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>USA</td>
<td>68</td>
<td>30</td>
</tr>
<tr>
<td>Canada</td>
<td>29</td>
<td>13</td>
</tr>
<tr>
<td>UK</td>
<td>27</td>
<td>12</td>
</tr>
<tr>
<td>The Netherlands</td>
<td>11</td>
<td>5</td>
</tr>
<tr>
<td>Japan</td>
<td>10</td>
<td>4</td>
</tr>
<tr>
<td>Germany</td>
<td>10</td>
<td>4</td>
</tr>
<tr>
<td>Switzerland</td>
<td>7</td>
<td>3</td>
</tr>
<tr>
<td>France</td>
<td>7</td>
<td>3</td>
</tr>
<tr>
<td>Italy</td>
<td>7</td>
<td>3</td>
</tr>
<tr>
<td>China</td>
<td>5</td>
<td>2</td>
</tr>
<tr>
<td>Australia</td>
<td>5</td>
<td>2</td>
</tr>
<tr>
<td>Singapore</td>
<td>4</td>
<td>2</td>
</tr>
<tr>
<td>Sweden</td>
<td>4</td>
<td>2</td>
</tr>
<tr>
<td>Thailand</td>
<td>4</td>
<td>2</td>
</tr>
<tr>
<td>Spain</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>South Korea</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>Europe</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>Belgium</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>Brazil</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>Norway</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>Taiwan</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>Mexico</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>Iran</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>Greece</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>Colombia</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>Ireland</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>Israel</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>New Zealand</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>Finland</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>Saudi Arabia</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>South Africa</td>
<td>1</td>
<td>0</td>
</tr>
</tbody>
</table>

\(^1\) This column sums to more than 222 as some studies were conducted in multiple countries
### Appendix 2.6: Focus of Economic Evaluations of Precision Medicine (2007-2017)

<table>
<thead>
<tr>
<th>Condition</th>
<th>Number of Studies (Total n=222)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Breast Cancer</td>
<td>104</td>
</tr>
<tr>
<td>Cardiovascular Conditions(^a)</td>
<td>22</td>
</tr>
<tr>
<td>Lung Cancer</td>
<td>18</td>
</tr>
<tr>
<td>Colorectal Cancer</td>
<td>15</td>
</tr>
<tr>
<td>HIV</td>
<td>12</td>
</tr>
<tr>
<td>Infectious Disease(^b)</td>
<td>7</td>
</tr>
<tr>
<td>Gastrointestinal Cancer</td>
<td>5</td>
</tr>
<tr>
<td>Epilepsy</td>
<td>4</td>
</tr>
<tr>
<td>Smoking Cessation</td>
<td>4</td>
</tr>
<tr>
<td>Rheumatoid Arthritis</td>
<td>3</td>
</tr>
<tr>
<td>Depression</td>
<td>3</td>
</tr>
<tr>
<td>Thyroid Nodules</td>
<td>3</td>
</tr>
<tr>
<td>Gout</td>
<td>2</td>
</tr>
<tr>
<td>Renal Disorders</td>
<td>2</td>
</tr>
<tr>
<td>Diabetes</td>
<td>2</td>
</tr>
<tr>
<td>Ovarian cancer</td>
<td>2</td>
</tr>
<tr>
<td>Endometrial Cancer</td>
<td>1</td>
</tr>
<tr>
<td>Contraception</td>
<td>1</td>
</tr>
<tr>
<td>Autoimmune Disease</td>
<td>1</td>
</tr>
<tr>
<td>Cystic Fibrosis</td>
<td>1</td>
</tr>
<tr>
<td>Non-diabetic Neuropathy</td>
<td>1</td>
</tr>
<tr>
<td>Pediatric Cancers</td>
<td>1</td>
</tr>
<tr>
<td>Leukemia</td>
<td>1</td>
</tr>
<tr>
<td>A range of cancers</td>
<td>1</td>
</tr>
<tr>
<td>Pulmonary Fibrosis</td>
<td>1</td>
</tr>
<tr>
<td>Familial Hypercholesterolemia</td>
<td>1</td>
</tr>
<tr>
<td>Mild Cognitive Impairment</td>
<td>1</td>
</tr>
<tr>
<td>Abdominal lesions</td>
<td>1</td>
</tr>
<tr>
<td>Malignant Melanoma</td>
<td>1</td>
</tr>
<tr>
<td>Prostate cancer</td>
<td>1</td>
</tr>
<tr>
<td>Head/Neck Cancers</td>
<td>1</td>
</tr>
</tbody>
</table>
Nasopharyngeal Cancers

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
</tr>
</tbody>
</table>

\[a \text{ including acute myocardial infarction, stroke prevention and acute coronary syndrome}
\]

\[b \text{ including Hepatitis C}
\]
## Appendix 2.7: Summary of Studies Which Discuss Capacity Constraints

<table>
<thead>
<tr>
<th>Author (Year)</th>
<th>Disease Area</th>
<th>Intervention and Comparator</th>
<th>How did the study discuss capacity?</th>
<th>Did the study quantify the impact of capacity constraints in the analysis? If so, what methods were used?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Barone et al., (2014) (141) Italy</td>
<td>Colorectal Cancer</td>
<td>Intervention: KRAS testing to guide treatment with cetuximab and bevacizumab Comparator: Standard chemotherapy</td>
<td>Used data that suggested 75% of experts thought that test turnaround time affects treatment choice. Turnaround longer than 15 days could limit therapy choice and 10 days is considered optimal. However, 25% of participants in a survey recording turnaround time stated a delay of over 15 days.</td>
<td>No</td>
</tr>
<tr>
<td>Breijer et al., (2012) (142) The Netherlands</td>
<td>Endometrial Cancer</td>
<td>Intervention: Three transvaginal sonography-based diagnostic strategies for endometrial cancer for women with post-menopausal bleeding. Comparator: Diagnosis based on patient history only</td>
<td>Recommend the use of patient characteristics to target diagnostics more effectively Highlighted the need to make the diagnostic stratifying algorithm easy to access and use for GPs due to small expected gains per patient</td>
<td>No</td>
</tr>
<tr>
<td>Chen and Dowdy (2014) (149) USA</td>
<td>HIV</td>
<td>Intervention: Risk calculators to guide HIV pre-exposure prophylaxis use in men who have sex with men Comparator: No pre-exposure prophylaxis</td>
<td>Economic evidence may be useful in guiding the implementation of interventions at a local and state level</td>
<td>No</td>
</tr>
<tr>
<td>Collinson et al., (2013) (137) UK</td>
<td>Acute Myocardial Infarction</td>
<td>Intervention: Blood based troponin assays and cardiac biomarkers for detecting AMI Comparator: Hospital admission and 10 to 12 hour troponin measurement</td>
<td>Made suggestions in a general observations section that test performance for new biomarkers should be adequate and that the marker should be measurable in a routine laboratory. Went on to say that measurement should be precise and accurate and there should be a rapid</td>
<td>No</td>
</tr>
<tr>
<td>Study Source</td>
<td>Cancer Type</td>
<td>Intervention</td>
<td>Comparator</td>
<td>Note Description</td>
</tr>
<tr>
<td>--------------------------------------------------</td>
<td>----------------------</td>
<td>---------------------------------------------------------------------------------------------------</td>
<td>--------------------------------------------------------------------------</td>
<td>----------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Contreras-Hernandez et al., (2008) (127) Mexico</td>
<td>Gastrointestinal</td>
<td>Second-line treatment with imatinib or sunitinib in patients with advanced gastrointestinal stromal tumours</td>
<td>Palliative care</td>
<td>Stated the need to take account of the fact that budgets of medical payers are constrained</td>
</tr>
<tr>
<td>Delea et al., (2012) (154) UK</td>
<td>Breast Cancer</td>
<td>Lapatinib and capecitabine for women with HER2 positive metastatic breast cancer who have previously received trastuzumab</td>
<td>Capecitabine monotherapy</td>
<td>Included a figure for drug wastage. Mentioned that the drug wastage parameter they included was similar to the value found in an Italian study prior to the implementation of a strategy to reduce waste.</td>
</tr>
<tr>
<td>Delea et al., (2013) (155) UK</td>
<td>Breast Cancer</td>
<td>Lapatinib and letrozole for post-menopausal women with hormone receptor and HER2 positive metastatic breast cancer</td>
<td>Trastuzumab and anastrozole or trastuzumab alone or letrozole alone</td>
<td>The authors noted that a lower rate of drug wastage after full programme implementation was observed</td>
</tr>
<tr>
<td>Dionne et al., (2012) (297) Canada</td>
<td>Paediatric Cancer</td>
<td>Genetic test to predict susceptibility to adverse events of cisplatin based chemotherapy</td>
<td>No test.</td>
<td>Suggested that test costs could change with higher utilisation and technological advances</td>
</tr>
<tr>
<td>Study</td>
<td>Disease</td>
<td>Intervention</td>
<td>comparator</td>
<td>Key points</td>
</tr>
<tr>
<td>------------------------------</td>
<td>------------------</td>
<td>------------------------------------------------------------------------------</td>
<td>-----------------------------------------------------------------------------</td>
<td>-----------------------------------------------------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Djalalov et al., (2014)</td>
<td>Lung Cancer</td>
<td>Intervention: EML4-ALK Fusion testing and first-line crizotinib treatment for patients with advanced ALK-positive non–small-cell lung cancer</td>
<td>Comparator: Platinum doublet chemotherapy</td>
<td>Suggested there were challenges to implementing molecular testing with FISH for ALK fusions: large number of NSCLC patients by low level of the mutation, testing; complexity of test; testing cost; need for sufficient tissue sample.</td>
</tr>
<tr>
<td>Essers et al., (2010)</td>
<td>Breast Cancer</td>
<td>Intervention: Trastuzumab for the adjuvant treatment of HER2 positive early breast cancer</td>
<td>Comparator: Observation and trastuzumab if progression to metastatic disease</td>
<td>Suggested that improving the transferability of economic evaluations may improve speed of reimbursement of new pharmaceuticals</td>
</tr>
<tr>
<td>Garrison and Veenstra</td>
<td>Breast Cancer</td>
<td>Intervention: Trastuzumab for various stages of HER2 positive breast cancer</td>
<td>Comparator: Unclear</td>
<td>Suggested that increasing the number of indications for a drug and, therefore, the number of eligible patients will change relative cost-effectiveness over time. However, the analysis still assumed that all patients immediately have access to the treatment when their indication was approved</td>
</tr>
<tr>
<td>Garrison et al., (2013)</td>
<td>Breast Cancer</td>
<td>Intervention: Retesting IHC tested HER2 negative patients for HER2 with FISH to reduce false-negative results</td>
<td>Comparator: Current testing algorithm (no retesting)</td>
<td>Suggested several factors that will affect the implementation of expanded reflex testing. Many labs only have the capacity to perform IHC and not FISH. Improvements in capacity and investment in equipment may be needed. However the requirements on increased capacity and equipment would be even greater if full initial HER2 testing were implemented.</td>
</tr>
<tr>
<td>Study</td>
<td>Disease</td>
<td>Intervention</td>
<td>Comparator</td>
<td>Conclusion</td>
</tr>
<tr>
<td>----------------------</td>
<td>----------</td>
<td>-------------------------------------------------------------------------------</td>
<td>-------------------------------------------------</td>
<td>---------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Hornberger et al., (2011) (145) USA</td>
<td>Breast Cancer</td>
<td>Intervention: 21-gene assay to guide treatment of patients with estrogen receptor positive, lymph node negative early stage breast cancer</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Comparator: Unclear</td>
<td>Suggested the relative cost-effectiveness of the assay may change with the volume of testing in clinical practice as a result of learning effects by clinicians using the assay</td>
<td>No</td>
</tr>
<tr>
<td>Ito et al., (2013) (130) USA</td>
<td>Breast Cancer</td>
<td>Intervention: Eliminating co-payments made by Medicare beneficiaries for aromatase inhibitors for patients with hormone-receptor positive early breast cancer</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Comparator: Existence of co-payments by Medicare beneficiaries for aromatase inhibitors for patients with hormone-receptor positive early breast cancer</td>
<td>Suggested that the drugs may be underused due to cost constraints but unclear whether from what was written whether this was from an organisational or patient perspective</td>
<td>No</td>
</tr>
<tr>
<td>Klang et al., (2010) (143) Israel</td>
<td>Breast Cancer</td>
<td>Intervention: Oncotype DX use in women with estrogen receptor positive, lymph-node negative early stage breast cancer</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Comparator: “Traditional prognostic pathways”</td>
<td>Suggested that clinicians would need to learn about the technology and how to interpret tests</td>
<td>No</td>
</tr>
<tr>
<td>Study Authors</td>
<td>Cancer Type</td>
<td>Intervention</td>
<td>Financial Implications</td>
<td>Conclusion</td>
</tr>
<tr>
<td>-----------------------</td>
<td>------------------</td>
<td>------------------------------------------------------------------------------</td>
<td>----------------------------------------------------------------------------------------</td>
<td>---------------------</td>
</tr>
<tr>
<td>Kondo et al., (2008)</td>
<td>Breast Cancer</td>
<td>Intervention: 21-gene reverse transcriptase-polymerase chain reaction assay in guiding treatment for lymph-node-negative, estrogen-receptor-positive, early-stage breast cancer</td>
<td>Suggested the need to consider the financial implications of needing an additional 12,000 tests a year and the need to do a budget impact analysis</td>
<td>Yes: used budget impact analysis to determine total cost implication</td>
</tr>
<tr>
<td></td>
<td>Japan</td>
<td>Comparator: National Comprehensive Cancer Network (NCCN) guideline/St Gallen recommendation-guided treatment</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lala et al., (2013)</td>
<td>Acute Coronary Syndrome</td>
<td>Intervention: CYP2C19 gene testing to guide antiplatelet therapy in patients with acute coronary syndrome undergoing percutaneous coronary intervention</td>
<td>Suggested that the time between performing the test and receiving test results is likely to shorten over time and with the advent of new technology making point of care testing more feasible</td>
<td>No</td>
</tr>
<tr>
<td></td>
<td>USA</td>
<td>Comparator: Two no-testing strategies</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lee et al., (2010)</td>
<td>Breast Cancer</td>
<td>Intervention: Adjuvant hormonal treatments for women with postmenopausal hormone-receptor positive early breast cancer</td>
<td>Suggested that the availability of treatment might be limited given the high cost of the drug and increasing healthcare budget constraints.</td>
<td>No</td>
</tr>
<tr>
<td></td>
<td>South Korea</td>
<td>Comparator: Tamoxifen</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lidgren et al., (2008b)</td>
<td>Breast Cancer</td>
<td>Intervention: HER2 testing and trastuzumab in combination with chemotherapy for metastatic breast cancer</td>
<td>Concluded that even if drug is a cost-effective use of resources that some clinicians may not prescribe it due to budget constraints. This would lead to inequitable access to treatment for patients</td>
<td>No</td>
</tr>
<tr>
<td></td>
<td>Sweden</td>
<td>Comparator: Chemotherapy alone</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lorenzana et al., (2012)</td>
<td>HIV</td>
<td>Intervention: Genotype assay for selection of third-line antiretroviral therapy</td>
<td>Suggested that tests costs are often cited as a barrier to implementation. The cost of the genotype test was found to have little impact</td>
<td>Yes – test cost was apparently varied in sensitivity analysis with suggestions that</td>
</tr>
<tr>
<td>Study</td>
<td>Cancer Type</td>
<td>Intervention</td>
<td>Comparator</td>
<td>Comment</td>
</tr>
<tr>
<td>-------------------------------------------</td>
<td>-------------</td>
<td>-----------------------------------------------------------------------------</td>
<td>---------------------------------------------------------------------------</td>
<td>-------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Machanda et al., (2015)(301) UK and USA</td>
<td>Breast</td>
<td>Intervention: Population based BRCA testing with varying Ashkenazi Jewish ancestry</td>
<td>Comparator: No BRCA testing</td>
<td>Suggested that to implement a population based testing strategy there will need to be wide-scale dissemination of information and knowledge involving working with stakeholders and health professionals. Other issues may be faced with “health system delivery, referral and management pathways, logistics, and control, which can vary across different models of care in different countries” No</td>
</tr>
<tr>
<td>McCowan et al., (2013)(153) UK</td>
<td>Breast</td>
<td>Intervention: High adherence to tamoxifen in women with breast cancer</td>
<td>Comparator: Low adherence (&lt;80%)</td>
<td>The issues raised in the paper focussed on adherence and as the authors state “besides patient-related factors, the characteristics of the disease and its treatment, the attributes of the health-care system and service delivery may also influence adherence”. Lower adherence was linked to shorter time to cancer recurrence and higher health service costs later on. Interventions which improve adherence should improve outcomes and may be highly cost-effective Yes – evaluation conducted across sub-groups of patients with under or over 80% adherence. Low adherence associated with expected loss of 1.12 discounted QALYs and increase of £5,970 in medical costs. Methods could be extrapolated to compare cost-effectiveness of high patient access to treatments.</td>
</tr>
<tr>
<td>Medical Advisory Secretariat (2010) (138)</td>
<td>Lung</td>
<td>Intervention: Epidermal growth factor Receptor mutation (EGFR) testing for prediction</td>
<td></td>
<td>Suggested there are barriers to routine FISH testing including lack of expertise in molecular techniques and lack of experience with No</td>
</tr>
<tr>
<td>Country/Location</td>
<td>Disease</td>
<td>Intervention</td>
<td>Comparator</td>
<td>Notes</td>
</tr>
<tr>
<td>------------------</td>
<td>---------</td>
<td>--------------</td>
<td>------------</td>
<td>-------</td>
</tr>
<tr>
<td>Canada</td>
<td>Breast Cancer</td>
<td>Gene expression profiling for guiding adjuvant chemotherapy decisions in women with early estrogen or progesterone receptor positive breast cancer</td>
<td>No testing</td>
<td>Suggested that the cost of testing meant that economic evaluation needed as a step before introducing into dynamic health systems. Also state that “clinician communication affects the usefulness of test information” for treatment decision making. Furthermore a lack of genetic literacy amongst patients may make it hard for them to fully take on information from genetic tests.</td>
</tr>
<tr>
<td>Patrick et al., (2009) (25) USA</td>
<td>Stroke Prevention</td>
<td>Genotype-guided (CYP2C9 and VKORC1) warfarin dosing for patients with atrial fibrillation</td>
<td>No testing</td>
<td>Suggested that the total cost of implementing genetic testing would have significant economic consequences</td>
</tr>
<tr>
<td>Retèl et al., (2010) (150) The Netherlands</td>
<td>Breast Cancer</td>
<td>70-gene signature MammaPrint for guiding treatment of patients with node negative breast cancer</td>
<td>St Gallen guidelines and Adjuvant online algorithm for guiding treatment of patients with node negative breast cancer</td>
<td>Stated that the analysis assumed perfect implementation and that the cost-effectiveness results represent the intervention being used in a perfect way.</td>
</tr>
<tr>
<td>Retèl et al., (2012) (136)</td>
<td>Breast Cancer</td>
<td>Potential scenarios for the intervention</td>
<td>The stated focus of the analysis in this paper is to Yes – the researchers</td>
<td>Yes – the researchers</td>
</tr>
<tr>
<td>Country</td>
<td>Cancer Type</td>
<td>Intervention</td>
<td>Comparator</td>
<td>Considered scenarios and barriers</td>
</tr>
<tr>
<td>--------------</td>
<td>-------------</td>
<td>--------------------------------------------------------------------------------</td>
<td>----------------------------------------------------------------------------</td>
<td>-----------------------------------</td>
</tr>
<tr>
<td>The Netherlands</td>
<td>-</td>
<td>Diffusion of a 70-gene signature into clinical practice</td>
<td>Comparator: Retaining the St Gallan guidelines and Adjuvant! Online algorithm for decision making</td>
<td>Consider the impact of partial implementation of a new technology. The analysis factored in a range of scenarios including barriers or removal of barriers. The scenarios evaluated were: hesitant adopters, increased user-friendliness, proof of the value of RNA-preservation, adoption in national guidelines of the test, improved reimbursement of testing, introduction of competitor tests, Oncotype dx (a rival test) is revealed to be better than the Mammaprint test, the test is made obsolete by a newer test, the test becomes available on the free market, and the better regulation of Mammaprint improves its market share versus Oncotype dx. The cost-effectiveness of Mammaprint over time and the diffusion of the intervention was explored by the researchers.</td>
</tr>
<tr>
<td>USA</td>
<td>Lung Cancer</td>
<td>Intervention: Multiplexed testing for EGFR and ALK mutations to guide NSCLC treatment</td>
<td>Comparator: No testing and treatment with pemetrexed and cisplatin</td>
<td>Accounted for the issue that long turnaround times for tests and inadequate tissue samples may push the balance of testing decisions towards generic chemotherapy use.</td>
</tr>
<tr>
<td>USA</td>
<td>Breast Cancer</td>
<td>Intervention: Population-based BRCA1/2 testing and ovarian cancer prevention for Ashkenazi Jews</td>
<td>Comparator: No testing</td>
<td>Suggested that improvements in other technologies such as MRI scans may cause additional costs and benefits for the intervention of interest when they are implemented. They also suggested that active implementation frameworks rather than passive uptake of guidelines may help the health system to achieve economies of</td>
</tr>
<tr>
<td>Study</td>
<td>Disease/Condition</td>
<td>Intervention</td>
<td>Comparator</td>
<td>Summary</td>
</tr>
<tr>
<td>--------------------------------------</td>
<td>-------------------</td>
<td>-------------------------------------------------------------------------------</td>
<td>------------</td>
<td>------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Saokaew et al., (2014) (126) Thailand</td>
<td>Gout</td>
<td>HLA-B5801 genotyping of gout patients with a high risk of allopurinol-induced severe cutaneous adverse reactions before treatment with allopurinol Comparator: No genetic testing</td>
<td>No</td>
<td>Stated that the analysis did not account for the feasibility of implementing the proposed test or take account of the affordability of the test. They also suggested that there is a need for clinicians to have an understanding of the test and its implications.</td>
</tr>
<tr>
<td>Segui et al., (2014) (125) Spain</td>
<td>Breast Cancer</td>
<td>70-gene signature to assess tumour recurrence risk and determine chemotherapy treatment Comparator: No testing</td>
<td>No</td>
<td>Suggested that the technical ease of use and pricing of the assay will be important when considering the implementation of tests into clinical practice.</td>
</tr>
<tr>
<td>Thompson et al., (2014) (15) UK</td>
<td>Autoimmune Diseases</td>
<td>Thiopurine-methyl transferase testing prior to prescription of azathioprine to predict susceptibility to adverse drug reactions Comparator: No testing</td>
<td>No</td>
<td>Suggested that the results observed in the prospective trial were the results not adhering to test recommendations regarding the dose of azathioprine, which infers there might be imperfect implementation of the test in clinical practice.</td>
</tr>
<tr>
<td>Vanderlaan et al., (2011) (152) USA</td>
<td>Breast Cancer</td>
<td>21-gene assay for women with early-stage, minimally node-positive, estrogen receptor–positive (N+(1-3)/ER+) HER2-negative breast cancer Comparator: No assay</td>
<td>No</td>
<td>A sensitivity analysis included variations in utilisation rates of testing. However, as the costs and benefits of testing were not dependent on the utilisation rate, the cost-effectiveness of the intervention was not affected by changes in the utilisation rate.</td>
</tr>
</tbody>
</table>

¹Title formatting has been changed to avoid clashes with the thesis formatting
Understanding barriers to the introduction of precision medicines in non-small cell lung cancer: A qualitative interview protocol [version 1; peer review: 2 approved]

Stuart Wright 1, Gavin Daker-White2, William Newman3, Katherine Payne1

1Manchester Centre for Health Economics, Division of Population Health, Health Services Research & Primary Care, University of Manchester, Manchester, M13 9PL, UK
2Centre for Primary Care, Division of Population Health, Health Services Research & Primary Care, University of Manchester, Manchester, M13 9PL, UK
3Manchester Centre for Genomic Medicine, University of Manchester, Manchester, M13 9PL, UK

Abstract

Background: While precision medicines targeting genetic mutations and alterations in non-small cell lung cancer (NSCLC) have been available since 2010, their adoption into clinical practice has been slow. Evidence suggests that a number of barriers, such as insufficient clinician knowledge, a need for training of test providers, or a lack of specific clinical guidelines, may slow the implementation of precision in general. However, little attention has been given to the barriers to providing precision medicines in NSCLC. The purpose of this protocol is to outline the design for a qualitative interview study to identify the barriers and facilitators to the provision of precision medicines for NSCLC.

Methods: This study will use semi-structured interviews with clinicians (n=10), test providers (n=10), and service commissioners (n=10) to identify the perceived barriers and facilitators to providing historical, current, and future precision medicines in NSCLC. Participants will be identified through mailing list advertisements and snowball sampling. Recruitment will continue until data saturation, indicated by no new themes arising from the data. Interviews will be conducted by telephone to facilitate geographical diversity. The qualitative data will be analysed using a framework analysis with themes anticipated to relate to: relevant barriers to providing precision medicines, the impact of different barriers on medicine provision, changes in the ability to provide precision medicines over time, and strategies to facilitate the provision of precision medicines.

Ethics: This study has been approved by the University of Manchester Proportionate Review Research Ethics Committee (Reference number: 2017-1885-3619). Written consent will be obtained from all participants.

Conclusion: This study is the first to explore the barriers and facilitators to providing precision medicines for NSCLC in the English NHS. The findings will inform strategies to improve the implementation of future precision medicines.
medicines. These findings will be disseminated in peer-reviewed publications and national and international conferences.

Keywords

Precision Medicine, stratified medicine, qualitative interviews, capacity, barriers, non-small cell lung cancer, cost effectiveness

Corresponding author: Stuart Wright (stuart.wright-2@manchester.ac.uk)

Author roles: Wright S: Conceptualization, Funding Acquisition, Methodology, Project Administration, Writing – Original Draft Preparation, Writing – Review & Editing; Daker-White G: Conceptualization, Funding Acquisition, Methodology, Supervision, Writing – Review & Editing; Newman W: Conceptualization, Funding Acquisition, Methodology, Supervision, Writing – Review & Editing; Payne K: Conceptualization, Funding Acquisition, Methodology, Project Administration, Supervision, Writing – Review & Editing

Competing interests: No competing interests were disclosed.

Grant information: This work was supported by a Wellcome Trust grant [203405] as part of a Society and Ethics Doctoral Studentship for the PhD programme of Stuart Wright

The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Copyright: © 2018 Wright S et al. This is an open access article distributed under the terms of the Creative Commons Attribution Licence, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

How to cite this article: Wright S, Daker-White G, Newman W and Payne K. Understanding barriers to the introduction of precision
Introduction

The concept of precision medicine is gaining increased attention as a potentially effective and cost-effective approach to the treatment of patients (Hatz et al., 2014; Payne et al., 2018; Phillips et al., 2014). Currently, the applied examples of precision medicine use a companion test-treat strategy to separate patients into groups according to the likelihood of responding to treatment or experiencing side effects. Medicines that use a companion test are now available for the management of non-small cell lung cancer (NSCLC) in clinical practice. The first example, gefitinib, was licensed in Europe in 2009, making it available for use in clinical practice. The medicine was then required to be made available to all eligible patients in the English NHS from 2010 with the approval of the drug as part of the National Institute for Health and Care Excellence (NICE) technology appraisal process (National Institute for Health and Care Excellence, 2010). Gefitinib was appraised by NICE to provide sufficient benefits given its costs for patients with advanced NSCLC who had cancer which tested positive for mutations which lead to overexpression of epidermal growth factor receptor (EGFR). Further treatments for EGFR positive tumours were approved for recommendation by NICE in 2012 (erlotinib) and 2014 (afatinib) (National Institute for Health and Care Excellence, 2014; National Institute for Health and Care Excellence, 2012). In 2012, crizotinib was licensed and subsequently was approved by NICE in 2014. This intervention involves targeting treatment using a test to detect a special type of mutation affecting the anaplastic lymphoma kinase (ALK) gene called a translocation whereby chromosomes break and re-join creating fusion genes with increased activity. In 2016 ceritinib, that is also targeted to ALK mutations, was also approved by NICE (National Institute for Health and Care Excellence, 2013; National Institute for Health and Care Excellence, 2016a). Treatments targeting the overexpression of programmed death ligand 1 (PD-L1) have been licenced and approved and treatments targeting BRAF gene mutations have a product license and are currently undergoing appraisal by NICE (National Institute for Health and Care Excellence, 2016b; National Institute for Health and Care Excellence, 2016c).

When treatments are recommended for use in the NHS by NICE, service commissioners are legally required to provide patients with access to these medicines within three months of the positive recommendation (NHSLA, 2015). Although the provision of companion diagnostics alongside precision medicines in cancer has recently become mandatory, historically this was not true (National Institute for Health and Clinical Excellence, 2011; The British in Vitro Diagnostics Association et al., 2016). Evidence produced by the charity Cancer Research UK has suggested that there is a significant lag in the provision of precision testing from when the precision medicine is first licensed to all patients gaining ready access to the medicine in the NHS (Cancer Research UK, 2015). The study, using a survey of 56 laboratories which were known to conduct molecular testing, estimated that in 2014, 48% of patients were not receiving mutation tests and this could mean that approximately 1,428 out of 3,007 patients who could have benefitted from EGFR targeting therapies were missing out. Prompt receipt of test results is required to inform clinical decision making but in some areas where testing was required to be made available for all patients, fewer than 50% of patients had EGFR test results available at their first consultation (Evans et al., 2013). Potential benefits from increasing the proportion of patients who receive a precision medicine have been improved life expectancy and fewer severe side effects than those on standard chemotherapies (Banz et al., 2011).

The availability of cancer somatic (tumour) mutation testing in the NHS has been repeatedly discussed during the appraisals of medicines using a test to direct treatment as part of the NICE appraisal process. In 2010, it was highlighted in the NICE appraisal of gefitinib that EGFR testing was not routinely conducted in the NHS, but it was said that the capacity to provide testing was present (National Institute for Health and Care Excellence, 2010). In 2012, the assessment report for erlotinib stated that testing had become standard practice (National Institute for Health and Care Excellence, 2012). However, the 2014 technology appraisal report for afatinib highlighted that there was still regional variation in the turnaround time for tests (National Institute for Health and Care Excellence, 2014). As such, there is evidence that not all patients had access to precision medicines for NSCLC, or their requisite tests over four years after approval, despite the requirement of access being provided after 3 months. As a result, patients were not receiving interventions which could have provided improved length and quality of life which should have been available to them.

Previous research

A study using face-to-face semi-structured interviews which explores how oncologists’ perceptions and work environment affect their use of genomic-targeted medicines is currently being undertaken in the United States (Chen et al., 2015). The published protocol for this study presents the results of a pilot study but it is difficult, however, to generalise outside the US setting that has a specific privately funded healthcare system. Under the remit of NHS England, which is funded by a tax-based healthcare system, all medicines recommended by NICE are legally required to be made available to patients (NHSLA, 2015). Furthermore, in the United States, there is no obligation to provide treatments and decision making will likely be more devolved to clinicians and patients rather than the more guideline focussed UK.

There may be some commonalities in clinician experiences between the US and UK. Approximately a third of the ten oncology fellows interviewed in the pilot study were uncertain about guidelines regarding the use of precision treatments as second or third line treatments for lung cancer while a third of those interviewed were also uncertain regarding how to order testing. Common barriers to performing tests included insufficient tissue samples, the inconvenience of testing and the cost of testing. Facilitators of tests were the ease of testing and deceiving results, as well as patients having health insurance. The cost of treatment was mentioned as a barrier by a smaller number of clinicians. These findings highlight how differences in financing arrangements may impact on the use of precision medicine in oncology. For beneficial targeted treatments to be prescribed, test results need to be available in a timely manner.
Delays in receiving test results were identified as a barrier to patients starting targeted therapies by half of the participants. These results mirror the findings of a US survey which found that the greatest perceived barrier to the use of precision medicines in practice was the cost of testing and targeted therapies (Petersen et al., 2014).

Other studies have sought to identify the barriers to precision medicine more generally. Taking account of this collective evidence base means the definition of precision medicine must also include tests for genetic predisposition for disease and tests for susceptibility to adverse events. In 2008, Newman and Payne identified that few clinical laboratories were offering pharma- macogenetic testing services (Newman & Payne, 2008). The timing of tests and coordination of testing with treatment was identified in qualitative interviews with stakeholders in breast cancer care as a key constraint of access to precision medicine alongside delays in testing whilst payer authorisation was sought (Weldon et al., 2012). In a 2013 study based in Canada, which used focus groups with physicians, the identified key relevant concerns about introducing precision medicine included: insufficient knowledge; a need for training of physicians; lack of specific guidelines and protocols for using tests; unequal access to testing due to socioeconomic differences; the financial burden of testing on public funds; additional time pressures that precision medicine will put on clinical practice; need for geneticist support after testing (Najafzadeh et al., 2013). The same authors also derived quantitative weights for the importance of different barriers in a subsequent study using a discrete choice experiment (Najafzadeh et al., 2012). The key attributes driving physicians’ preferences for using precision medicine were the availability of training and guidelines. Interestingly, this preference study also found two sub-groups with different types of preferences: one much more sensitive to the cost of the test than the other (Najafzadeh et al., 2012). Physicians in this group were more likely to be female.

In a systematic review of previous literature, with a particular focus on strategic reports from the European Commission funded PerMed – FP7 project, Horgan et al. (2014) identified a wide range of constraints to introducing precision medicine in Europe (Horgan et al., 2014). Barriers included limited resources, test turnaround time, lack of health professional knowledge and communication. Furthermore, the authors also identified barriers relevant to patients including a lack of awareness and under-standing of precision medicine and poor health literacy. Another significant barrier to implementing precision medicines across Europe will be understanding how reimbursement decisions can be made about such interventions given their unique properties (Payne & Annemans, 2013).

Despite the number of studies investigating barriers to the uptake of precision medicine in general, there has been a paucity of research focussing on the delayed implementation of interventions targeting NSCLC. Furthermore, there have been no studies examining the barriers to implementing precision medicine in the context of treating NSCLC in the UK NHS. As the number of precision medicines approved for use in NSCLC continues to expand, it becomes increasingly urgent to understand how best to implement such interventions in order to ensure that all relevant patients have access to potentially life extending and improving treatments.

Aim

The primary aim of this study is to explore the type and extent of barriers experienced by service providers and service commissioners when introducing licensed precision medicines for the treatment of NSCLC in relevant patient populations and individuals. Furthermore, a secondary aim is to identify strategies which have facilitated the improved provision of precision medicines for NSCLC in the English NHS.

Objectives

This study has four objectives:

• To identify the types of perceived clinical and organisational barriers to providing licensed test-treat medicines indicated for the treatment of NSCLC to patients;
• To explore the potential impact of the identified different barriers to the provision of licensed test-treat medicines indicated for the treatment of NSCLC;
• To explore how the availability of existing licensed test-treat medicines indicated for the treatment of NSCLC has changed over time;
• To identify strategies which have been used to improve the availability of licensed test-treat medicines indicated for the treatment of NSCLC.

Methods

This study will use semi-structured telephone interviews with clinicians, test providers, and service commissioners to identify the barriers to implementing licensed test-treat medicines indicated for the treatment of NSCLC. Previous research has shown that there have been issues with implementing precision medicine for lung cancer into the NHS but few have explored why this was the case for these medicines in particular (Chen et al., 2015). While quantitative analyses can assist in showing the number of patients not prescribed precision medicines it is useful to use qualitative methods to explore the reasons for this observation. Qualitative methods, such as semi-structured interviews, can be used to explore the thoughts, attitudes and opinions of those who were involved in implementing licensed test-treat medicines indicated for the treatment of NSCLC in clinical practice. In this study, semi-structured interviews will be used to understand the barriers to introducing licensed test-treat medicines indicated for the treatment of NSCLC. This approach also has the advantage of allowing the investigation of the perceived barriers of precision medicines introduced at different points in time by identifying the experiences and opinions of key stakeholders. This will allow the exploration of the potential for the English NHS to learn from previous implementation issues to improve future treatment provision.
Sampling
The sampling frame will aim to identify stakeholders with experience of providing and introducing licensed test-treat medicines indicated for the treatment of NSCLC. While demand side factors, such as uptake of treatment or adherence to medicines, linked to patients’ preferences for treatment may also impede the implementation of precision medicine, the focus of these interviews is to identify the supply side capacity constraints. Therefore, patients will not be interviewed in this study. The relevant stakeholders will be drawn from two groups: service providers, for example clinicians, pathologists, and geneticists; and service commissioners which may include individuals who are members of care commissioning groups or involved in commissioning at the national level through NHS England. The principle service providers of interest are oncologists and respiratory physicians specialising in lung cancer, but also geneticists and pathologists who are key in providing examples of tests used in licensed test-treat medicines indicated for the treatment of NSCLC, such as EGFR and ALK testing and the emerging PD-L1 test.

Purposive sampling will be used to gain a diverse sample in terms of the setting and geographical location of testing and treatment (Palinkas et al., 2015). This characteristic is likely to be important in the context of introducing precision medicines as experiences may vary depending on the size and nature of hospitals and trusts. For example, mutation testing services may be more readily available in larger teaching hospitals with established links to laboratories. For smaller, rural hospitals there may be a greater logistical challenge in sending samples for testing and receiving results in a timely manner.

Service providers with over 7 years of NHS experience will be targeted as such individuals are likely to have direct experience of the introduction of EGFR and ALK testing and treatment as they were working in clinical practice.

The service commissioner sample will comprise hospital, regional and national level individuals involved with service commissioning and funding decisions. Examples of service commissioners may involve members of care commissioning groups, hospital finance staff and decision makers involved with national organisations, such as the National Institute for Health and Care Excellence (NICE). As in recruitment for the clinician sample, geographical diversity will be sought through purposive sampling and service commissioners will be required to have been in a relevant position when EGFR and ALK mutation based testing and treatment were introduced.

Sample size
There are no defined rules for calculating sample size in qualitative studies (Patton, 2002). In quantitative studies, a sufficient sample size is required to identify statistically significant differences in the variables of interest. However, qualitative interviews are aimed at identifying the breadth of experiences, thoughts, or opinions on a given subject. This study will therefore start with an approximate sample of 10 clinicians or test providers, and 10 service commissioners but sampling will continue iteratively until no new themes are arising from the collected data, otherwise known as inductive thematic saturation (Saunders et al., 2017).

Recruitment
The clinician and test provider samples (n=20) will be recruited via the British Thoracic Oncology Group (BTOG) (British Thoracic Oncology Group, 2017) and the Royal College of Pathologists (RCPath) (Royal College of Pathologists, 2017). Details about the study and an invitation to participate will be circulated via the BTOG mailing list which currently has 2083 members and the RCPath that has over 11,000 members. Information regarding the study will be sent to participants using mailing lists, with contact details of the principal investigator provided for those interested in taking part. These individuals will then be sent more detailed information about the study.

Service commissioners (n=10) will be recruited using existing links and collaborations within the research team to identify an initial sample. Service commissioners will be directly sent an email including information about the study and the contact details of the principal investigator. Snowball sampling will be used for both samples whereby participants will be asked if they know any other individuals who meet the inclusion criteria who may be interested in taking part in the study (Lewis-Beck et al., 2004).

Telephone interviews
A bespoke telephone interview schedule has been created to address the key research questions while remaining open enough to allow relevant new lines of enquiry to be explored. Due to the focus of this work on capturing a geographically diverse sample to represent heterogeneity in health care provision, telephone interviews will be used to collect qualitative data (Musselwhite et al., 2007; Novick, 2008). Semi-structured interview schedules will be created for each sample, informed by a review of previous economic evaluations of precision medicines (including health technology assessments) and consultation with two expert clinical advisors who are lung oncologists and a patient representative group (Roy Castle Lung Cancer Foundation). An initial draft interview schedule for the clinician sample is presented in Supplementary File 1. While the core questions for each interview schedule will be similar, there will be slight variations in the way questions are asked depending on the particular role of the interviewee. For example, clinicians will be asked primarily about their experience offering treatments to patients while for geneticists and pathologists the focus will be on offering testing. Interviews are expected to last approximately 1 hour.

All interviews will be audio-recorded and transcribed verbatim by an approved, contracted transcription company (Associated Verbatim Reporters). Recordings will be sent via an encrypted data transfer.

Data analysis
The aim of the data analysis is to identify the range of barriers which may prevent patients’ access to precision treatments for NSCLC, to determine which are the most important barriers, and to identify strategies to improve the implementation of
The qualitative data will be analysed using a framework analysis facilitated by using the NVivo software (QSR International, 2017).

Framework analysis is a five stage process involving: familiarisation, identifying a thematic framework, indexing, charting, mapping and interpretation (Ritchie & Spencer, 1994). In the initial familiarisation stage, the researcher reads an initial set of the interviews in order to gain an understanding of the initial themes emerging from the data.

The initial key themes identified during the data familiarisation stage, alongside evidence from previous research, form an initial thematic framework against which the selection of data is sorted and collected (Gale et al., 2013). As semi-structured interviews are being used for this study, it is anticipated that many of the themes will originate in the questions contained in the interview schedule. As new themes emerge from the data, they are added to the framework. Each transcript is then indexed against these themes, with sections from the text which support different themes annotated for later retrieval. Charting brings the separate transcripts together to create a picture of the research as a whole (Ritchie & Spencer, 1994). A chart is drawn up featuring the identified themes and potentially sub-headings for these themes. Information from each participant’s transcript which links to these themes is recorded in the chart, keeping the order of participants the same in each theme.

This analysis of qualitative evidence in a systematic way facilitates the discovery of patterns in the data while highlighting deviant cases for further investigation. In the context of this study, this will be the range of barriers which occur in providing and accessing test-treat medicines for NSCLC and views about which barriers were most significant in restricting the provision of precision medicines. However, the use of charting will also make clear the types of respondent referring to different topics which will serve to highlight the different perspectives of the availability of precision medicines.

**Data storage and anonymisation**

Phone calls and recording will take place while the researcher is at The University of Manchester in an enclosed office. The recording device and memory card containing the interview recordings will be stored in a locked drawer in a secure university office. The recordings will be saved onto an encrypted university computer, and the files password protected. Transcription of interviews will be conducted by a university approved company with secure file transfer protocols. Recorded interviews will be deleted from recording devices after they have been stored on a computer and anonymised and then destroyed completely at the end of the study. Interview transcripts will be stored for 10 years.

Anonymisation will be achieved by removing references to participants’ names as well as any reference to information which could lead to identification of the participant such as the name of their place of work. When referencing data from the transcripts, generalised information regarding the participant will be provided to demonstrate their demographics whilst not allowing identification.

**Ethical concerns**

This study has been approved by the University of Manchester Proportionate Review Research Ethics Committee (Reference number: 2017-1885-3619).

Clinicians and service commissioners who are interested in taking part in the study will be asked in the mailing list adverts to email or phone the research team to express an interest in taking part. The researchers will then email the potential participant a participant information sheet (Supplementary File 2). After receiving an information sheet (Supplementary File 2), potential participants will be given at least 24 hours to consider taking part in the study. If they agree to take part they will be asked to complete a written consent form and to return a copy to the researchers by post or email (Supplementary File 3).

**Dissemination of findings**

This study will form part of the PhD thesis for Stuart Wright. It is anticipated that this will be submitted in September 2019. All research participants will be emailed a summary of the main findings of this study after data analysis has been completed. The research team also plan to publish the full study in a peer-reviewed journal and present the results at relevant national and international conferences, for example the annual BTOG conference. In line with the ethical approval for this project, the raw qualitative data will not be made publically available.

**Remuneration**

Participants will not receive remuneration for taking part in this interview study.

**Contribution of this study**

The aims of this study are to identify barriers to providing precision medicines to patients and strategies which may be used to improve patient access to the medicines. As precision medicine is a rapidly expanding area and NICE continues to evaluate new examples of precision medicine in NSCLC, it is hoped that by learning from previous examples of slow implementation of such interventions, the rate at which new interventions are incorporated into the healthcare service can be improved. This will ensure that patients have access to new treatments which offer the potential to improve their quality and quantity of life.

In addition to these broad aims, the results of this study will be directly used in the lead author’s PhD to inform economic models of example precision medicines in NSCLC. Currently economic evaluations conducted to determine the cost-effectiveness of such interventions do not take account of the barriers to their provision in the health service. It is feasible that if the costs and benefits of providing these interventions to patients are depend-ent on the level of implementation, then the cost-effectiveness of such treatments will also depend on their level of use in...
practice. It is therefore important to understand the barriers to using apparently cost-effective precision medicines to ensure that these are implemented in a manner which makes best use of the limited resources available in the health system. The lead author’s PhD will investigate how including these barriers impacts on cost-effectiveness estimates and the barriers identified in this study will be used in applied examples from NSCLC.

To date two interviews have been conducted for the study and recruitment is open for the clinician and service commissioner samples. Participants are being sought for pilot interviews in the test provider sample.

Data availability

All data underlying the results are available as part of the article and no additional source data are required.

Supplementary material

Supplementary File 1: Draft of the semi-structured telephone interview schedule Click here to access the data.

Supplementary File 2: Participant Information Sheet. Click here to access the data.

Supplementary File 3: Consent form for participants. Click here to access the data.

References


Competing interests

No competing interests were disclosed.

Grant information

This work was supported by a Wellcome Trust grant [203405] as part of a Society and Ethics Doctoral Studentship for the PhD programme of Stuart Wright.

The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Acknowledgements

The authors would like to thank Professor Fiona Blackhall and Dr Yvonne Summers for their comments on this protocol. We would also like to thank the participants who took part in the pilot interviews for this study.
Wright et al. present an interesting study protocol that aim to investigate the barriers experienced by service providers (clinicians and test providers) and service commissioners towards introduction of licensed and approved precision medicines for treatment of NSCLC. The study investigates the past and present experiences, as well as looking into potential future situations. The findings will provide important input to the planned upcoming studies on how including the identified barriers impacts the cost-effectiveness estimates (to be carried out in the context of the PhD of the lead author).

The rationale for and objectives of the study are clearly described. The perspective taken is also clear: focus is on the supply side constraints, rather than patient experiences.

The study design is appropriate for the aim of the study, as it involves interviewing with key stakeholders on the service provision and service commissioning sides. Nevertheless, there is a potential area of improvement, which is including a stakeholder analysis in this study, or adjacent to it.

Barriers involved in implementation of services in health systems do not only involve technical constraints. They are usually rooted in how the health system, including services, is organized. Stakeholders play a crucial role in how current status of service provision is shaped, and if and how new services can be introduced.

At this point, using stakeholder analysis to understand the power and interest of different stakeholders in introduction of precision medicines in treatment of NSCLC can provide valuable information. The stakeholder map including the power and interest grid can potentially provide the landscape and help to understand why the identified barriers exist. Based on this information, strategies can be developed to overcome the barriers, as this is also a part of the study aims.

The stakeholder analysis can be done retrospectively, i.e. to understand the status in past cases of EGFR and ALK and how this influenced their introduction into services, and also to understand the current (and potential future) cases. Stakeholder analysis is ideally carried out in the initiation and phase. But at this
moment, it can also be incorporated to the existing protocol with additional questions in the topic guide for identification of stakeholders and their potential power and interests.

The authors may think that they are well aware of the stakeholder landscape in English NHS and are already considering these in the study. Nevertheless, it should be considered that a structured analysis on stakeholders has the potential to provide further data and insights, enhancing their analysis approach (framework analysis). A stakeholder analysis may also provide data partially on the issues raised by the other reviewer, Brett Doble, such as how the pathology labs, as stakeholders, are organized.

The following article provides a good overview of stakeholder analysis method, involving mainly the power and interest of the stakeholders: Bryson, J. M. (2004). What to do when stakeholders matter: stakeholder identification and analysis techniques. Public management review, 6(1), 21-53. Doi: 10.1080/14719030410001675722

More detailed guides, involving stakeholders’ knowledge, position, alliances and resources on the matter, in addition to power and interest, can also be used (e.g. the WHO’s Stakeholder Analysis Guidelines).

The study protocol provides a detailed description of methods. I agree with the other reviewer’s suggestions for elaboration of potential biases, as well as strategies to limit the effects of them.

Is the rationale for, and objectives of, the study clearly described?
Yes

Is the study design appropriate for the research question?
Partly

Are sufficient details of the methods provided to allow replication by others?
Yes

Are the datasets clearly presented in a useable and accessible format?
Not applicable

Competing Interests: No competing interests were disclosed.

Reviewer Expertise: Integration of innovations into health systems; personalized medicine; personalized health care

I have read this submission. I believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.

Reviewer Report 19 March 2018

https://doi.org/10.21956/wellcomeopenres.15191.r31712

©2018 Doble B. This is an open access peer review report distributed under the terms of the Creative Commons Attribution Licence, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.
Brett Doble

Health Economics Research Centre, Nuffield Department of Population Health, University of Oxford, Oxford, UK

Wright et al. propose a very interesting study that looks to understand past, present and future barriers to the clinical implementation of what they term ‘precision medicine’ in non-small cell lung cancer (NSCLC) within the English National Health Service (NHS). To achieve this, they will use a number of semi-structured qualitative interviews with clinicians, test providers and service commissioners who will have likely been previously exposed to the challenges in implementing such approaches to care.

Overall, the protocol clearly articulates the main aims of the research and details appropriate methods to answer the research question. I do, however, suggest a few issues that the authors might wish to consider to improve the study as outlined below:

The authors have chosen to specifically focus on implementation issues for ‘precision medicine’ in NSCLC and have been quite prescriptive in the justification for their study (i.e., ‘no such evidence exists for this type of cancer). While I agree that this type of cancer provides an excellent case study to investigate such issues I wonder to what extent implementation issues are cancer-type specific and not just largely applicable to ‘precision medicine’ in all cancer types more generally. For example, I imagine that similar implementation issues were experienced when HER2 testing and trastuzumab treatment became available and that this may have already been the focus of considerable research. I do recognise that existing research might not be within the English context, but there may still be lessons that can be learned from such research to further focus your questions within the proposed qualitative interviews.

I suggest that the authors attempt to highlight why implementation issues in NSCLC are potentially different from other types of cancer. For example, is it because multiple test and targeted treatments are available at first-line? Or potentially the fact that most NSCLC patients present to the clinic with late stage disease compared to other cancer types, which may affect implementation issues for testing? It would be helpful if the authors indicate more clearly how their research has been designed to specifically target discussion of these unique issues rather than just obtain the same discussion on standard challenges that have been previously identified when implementing precision medicine in general.

Further to my point above, after reviewing the interview schedule it seems the questions being asked of the participants are very focused on the potentially outdated concept of single mutation tests and associated targeted treatments. One of the most unique things about precision medicine in NSCLC is the fact that multiplex testing potentially offers value in a first-line setting. This is highlighted in the authors protocol as they state that multiple genomic alterations are available to potentially guide first-line treatment (e.g., erlotinib and crizotinib). Given limited tumour tissue from biopsies it might be necessary that EGFR and ALK testing occur at the same time using a single test. I recognise that the UK might not be at this stage yet, but I think it might be useful to introduce a question/prompt concerning multiplex testing if this information does not come out in the questions concerning implementation of newer precision medicines.

I think the previous research section of the protocol could be improved to be more comprehensive of the entire evidence base. I would specifically look for evidence concerning implementation issues with HER2 testing and trastuzumab as I mentioned above as well as the fact that there has also been some research on implementation of genome sequencing in clinical practice that may be very closely related to the issues the authors are trying to identify. Furthermore, I would also suggest that the authors attempt to look at research conducted by sociologists in this area as I identified a number of relevant studies with just a
quick search that contain relevant insights, which may be used to better focus questions during the proposed qualitative interviews:

Samuel, G.N. and Farsides B. The UK’s 100,000 Genomes Project: manifesting policymakers’ expectations. New Genetics and Society. 2017


In addition, I wonder if the authors have considered taking a broader approach to understanding barriers to implementation by conducting ethnographic research, in which the interviews would form just one aspect of the data collection. I know this might be beyond the scope of this initial study, but it might be something to considering moving forward in order to put the results of the interviews into better context. One of the biggest implementation issues in the English NHS is likely to be technical issues in how tumour samples are collected and stored. This issue is further compounded by the fact that testing technology as well as knowledge concerning the genomics of cancer can rapidly change and that there are a number of different tests available for tumour genomic testing in lung cancer. To ensure uptake of testing, the NHS will require a huge organisational, professional and cultural shift as trying to change how tissue is handled is a core issue in how pathology departments work in the NHS (e.g., genomic tests in cancer require rapid turnaround time, which traditionally has not been required). Fundamentally, it comes down to whether or not pathology labs share the same vision with policy makers and are they likely to oppose changes to their current model of pathology testing. This of course is also compounded by the lack of resources within the NHS to accommodate transformational change. To get at these issues you could potentially conduct observations at different testing site with England. The pathway from obtaining a biopsy sample to the sample being prepared and processed through testing and final generation of the results could be directly observed and details of the process recorded. Such methods would enable a more complete interpretation of implementation issues as the interview material could be extended by observations of the testing pathway at different sites, while fieldwork notes could be elaborated in light of individual views and explanations obtained in the interviews. This is just a suggestion as to how you could further build on the research proposed in the protocol.

Finally, I suggest that the authors add some discussion in the protocol with regards to potential bias that may affect their results. For example, it would be helpful if the authors could comment on how they might look to address self-selection bias in their participant recruitment strategy, given participants most interested in this topic will be most likely to agree to participate in the study, particularly for the clinician sample. Will the authors make any attempt to engage with stakeholders that are not so keen on implementing precision medicine in NSCLC? Furthermore, on review of the interview schedule I noticed that some of the questions are asking participants to recall events that happened close to 10 years ago. Recall bias is likely to be an issue here, do the authors have any strategies to limit the effect on this on results (e.g., potentially through the use of prompts)?

Is the rationale for, and objectives of, the study clearly described?
Partly

Is the study design appropriate for the research question?
Partly

Are sufficient details of the methods provided to allow replication by others?
Yes
Are the datasets clearly presented in a useable and accessible format?
Not applicable

Competing Interests: No competing interests were disclosed.

I have read this submission. I believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.
Appendix 4.2: Participant Information Sheet

Understanding Barriers to the Introduction of Stratified Medicines in Lung Cancer

Participant Information Sheet

You are being invited to take part in a research study being conducted as part of a PhD project seeking to understand how economic evaluations of stratified medicines can incorporate capacity constraints. Before you decide whether to take part, it is important for you to understand why the research is being done and what it will involve. Please take time to read the following information carefully and discuss it with others if you wish. Please ask if there is anything that is not clear or if you would like more information. Take time to decide whether or not you wish to take part. Thank you for taking the time to read this.

Who will conduct the research?

This research is being conducted by Mr Stuart Wright MSc, Manchester Centre for Health Economics, as part of his PhD programme of study.

What is the purpose of the research?

This study is focussing on the introduction of stratified medicines for the treatment of non-small cell lung cancer (NSCLC). While such treatments have been approved for use in the National Health Service since 2010 (epidermal growth factor receptor tyrosine kinase inhibitors) and 2014 (anaplastic lymphoma kinase inhibitors), their use in clinical practice was slow to develop. This meant that many patients who may have benefitted from these treatments did not have access to them. The aim of this research is to identify the barriers that slowed the introduction of stratified treatments for non-small cell lung cancer. Identifying these barriers may inform the implementation of new stratified medicines in non-small cell lung cancer such as programmed death ligand 1 and B-Raf inhibitors.

Why have I been chosen?

You have been chosen as you are either a clinician or service commissioner involved in the provision or commissioning of stratified medicines in non-small cell lung cancer.

What would I be asked to do if I took part?

A researcher (Stuart Wright), will interview you by telephone about your experience in providing stratified medicines for non-small cell lung cancer. The interview will last approximately 1 hour and will be audio-recorded. However, your details will remain anonymous at all times.

In the interview we would like to hear about:
1. Your experience of the introduction of testing and treatments for patients with EGFR and/or ALK mutated non-small cell lung cancer.
2. The barriers that you experienced in providing these stratified treatments
3. What strategies you used to overcome the barriers to using stratified treatments

What happens to the data collected?

The data from the combined interview transcripts in this study will be used to address a number of objectives:
1. To identify the range of clinical and organisational barriers to patients receiving approved and recommended stratified testing and treatments for NSCLC.
2. To explore the relative impact of different barriers on patient access to stratified testing and treatments for NSCLC.
3. To explore how the availability of existing approved and recommended stratified testing and treatments for NSCLC developed over time.
4. To identify strategies which have been used to improve the availability of stratified testing and treatments for NSCLC over time.

This will be achieved by using a qualitative analysis technique known as framework analysis. Data is analysed as it is collected and early themes in the interviews are identified. These key themes, alongside evidence from previous research, form an initial thematic framework against which the selection of data is sorted and collected. As new themes emerge from the data, they are added to the framework.

In the context of this study, these themes will be the range of barriers which occur in providing and accessing stratified medicines in NSCLC as well as views about the importance of each issue. It is anticipated that the results of this study will help to inform the introduction of future stratified medicines by using learning from previous treatments to avoid future barriers.

How is anonymity maintained?

Identifiable data will be stored on university computers to allow the researcher to link participants responses to their demographic information. A private transcription company will be used to transcribe data from recordings to text. The recordings will be sent using encrypted transfers and no participant identifiable data will be provided to the company. Furthermore, in reporting and publishing the results of the study, no participant identifying information will be included in the document.

Study data and material may be looked at by individuals from the University of Manchester, from regulatory authorities or from the NHS Trust, for monitoring and auditing purposes and this may well include access to personal information.

Data will be archived according to the University of Manchester's policy which is currently 5 years. Consent forms will be retained as essential documents, but items such as contact details will be deleted as soon as they are no longer required.

What happens if I do not want to take part or if I change my mind?

It is up to you to decide whether or not to take part. If you do decide to take part you will be given this information sheet to keep and be asked to sign a consent form. If you decide to take part you are still free to withdraw up to the time of publication of the study in a peer reviewed journal or the student’s final thesis, without giving a reason and without detriment to yourself.

Will I be paid for participating in the research?

You will not be paid for taking part in this study.

What is the duration of the research?

This study will involve 1 telephone interview which will be approximately 1 hour in length.

Where will the research be conducted?

Interviews will be conducted by telephone. The researcher will be located at the University of Manchester and will telephone from a private office.
Will the outcomes of the research be published?

The outcomes of this research will constitute a part of the student’s PhD thesis. It is anticipated that this will be submitted in September 2019. Furthermore, the researchers will seek to publish the results of this research in a peer-reviewed journal.

Who has reviewed the research project?

This research has been reviewed by the researchers supervisors (Professor Katherine Payne and Dr Gavin Daker-White) and an overview of the research has been reviewed by an independent member of the student’s centre (Mr Niall Davison).

This project has been reviewed by the University of Manchester Proportionate Research Ethics Committee.

What if something goes wrong?

You are free to withdraw from the study at any time. No reasons need to be given. Any taped or paper record of the interview and your contact details will be destroyed.

If you are worried about any part of this study please contact the research team. You can phone Stuart Wright on 0161 306 7970 or Professor Katherine Payne on 0161 306 7906.

What if I want to make a complaint?

Minor complaints

If you have a minor complaint then you need to contact the researcher(s) in the first instance. We will respond to any complaints about the study. To complain you can telephone Professor Katherine Payne on 0161 306 7906 or email Katherine.payne@manchester.ac.uk.

Formal Complaints

If you wish to make a formal complaint or if you are not satisfied with the response you have gained from the researchers in the first instance then please contact the Research Governance and Integrity Manager, Research Office, Christie Building, University of Manchester, Oxford Road, Manchester, M13 9PL, by emailing: research.complaints@manchester.ac.uk or by telephoning 0161 275 2674 or 275 2046.

What Do I Do Now?

If you have any queries about the study or if you are interested in taking part then please contact the researchers:

STUART WRIGHT
EMAIL: stuart.wright-2@manchester.ac.uk
TELEPHONE: 0161 306 7970

This Project Has Been Approved by the University of Manchester’s Research Ethics Committee [Reference: 2017-1885-3619].
Appendix 4.3: Consent Form

Staff Interview Consent Form

Title of project: Understanding Barriers to the Introduction of Stratified Medicines in Lung Cancer

<table>
<thead>
<tr>
<th>1. I confirm that I have read the attached information sheet on the above project and have had the opportunity to consider the information and ask questions and had these answered satisfactorily.</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>2. I understand that my participation in the study is voluntary and that I am free to withdraw:</th>
</tr>
</thead>
<tbody>
<tr>
<td>* at any time</td>
</tr>
<tr>
<td>* without having to give a reason for withdrawing</td>
</tr>
<tr>
<td>* and without detriment to myself</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>3. I understand that the interview will be audio recorded.</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>4. I agree to the use of anonymous quotations from this interview in research outputs arising from the study</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>5. I understand that data collected during the study may be looked at by individuals from the University of Manchester, from regulatory authorities or from the NHS Trust, where it is relevant. I give permission for these individuals to have access to my records.</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>6. I agree to take part in this study</th>
</tr>
</thead>
</table>

Name of participant: ………………… Signed: ……………………… Date: …………………

Name of researcher: ………………… Signed: ……………………… Date: …………………

This project has been approved by the University of Manchester Research Ethics Committee
Appendix 4.4: Interview Schedule for Service Providers

Draft - Interview Schedule - Clinicians

(Phrases in brackets and bold represent prompts for the interviewer)

Thank you for agreeing to be interviewed.

I am a PhD student based in MCHE at The University of Manchester. I am funded by the Wellcome Trust

I want to ask you some questions to explore your experience with using stratified medicines to treat NSCLC.

The interview should take no longer than one hour and will be recorded. Firstly could you confirm that you have received and returned the consent form for this interview? Do you have any questions before we start?

General

How would you define precision medicine in the context of NSCLC?

Given this definition, could you list all of the currently licensed stratified medicines for NSCLC which you are aware of?

Probe: are you aware of any new stratified medicines currently being trialled

Which of these have you used in clinical practice? (link to questions below)

Experience of EGFR mutation testing and treatment

1. Were you working in this or a similar role when EGFR targeting interventions were introduced into the NHS (~2010)?

2. Could you tell me about your memory of what happened when EGFR was first introduced and subsequent events?

   Can you remember when you first prescribed an EGFR mutation targeting treatment?

   - (Were you aware of EGFR testing/treatments before they were approved? How did you hear of them?) (For example clinical trials, conferences etc)
   - (How did you hear of them if you were not aware of these interventions before approval?)
   - (Could you offer these interventions immediately after approval? What proportion of patients could receive these treatments?)
   - (What factors prevented you from offering testing and treatment to all patients?) (Probe but do not lead with previously identified barriers from literature)
   - (Which were the biggest constraints to offering and using EGFR based testing and treatments?)

3. Do you currently prescribe treatments that require EGFR mutation testing?
4. Are you aware of any guidelines regarding *EGFR* mutation testing in the NHS? (i.e. *national, regional hospital guidelines*)

4b. Are you aware of any guidelines regarding *EGFR* mutation based treatment in the NHS? (i.e. *national, regional hospital guidelines*)

5. In your experience, do all patients who are currently eligible to receive *EGFR* mutation testing and treatment have access to these interventions?
   - (What factors prevent you from offering testing and treatment to all patients?) (probe but do not lead with previously identified barriers from literature)
   - (Which are the biggest constraints to offering and using *EGFR* based testing and treatments?)

6. Are you aware of any strategies to increase the number of patients receiving *EGFR* mutation targeting treatments?

7. Are there any strategies which you believe could be implemented to further increase the number of patients receiving *EGFR* targeting treatments in the future?

**Experience of ALK mutation testing and treatment**

1. Were you working in this or a similar role when *ALK* targeting interventions were introduced into the NHS (~2010)?

   2. Could you tell me about your memory of what happened when *ALK* was first introduced and subsequent events?
      - Can you remember when you first prescribed an *ALK* mutation targeting treatment? (see prompts for *EGFR*)

3. Do you currently prescribe treatments that require *ALK* mutation testing?

4. Are you aware of any guidelines regarding *ALK* mutation testing in the NHS? (i.e. *national, regional hospital guidelines*)

4b. Are you aware of any guidelines regarding *ALK* mutation based treatment in the NHS? (i.e. *national, regional hospital guidelines*)

5. In your experience, do all patients who are currently eligible to receive *ALK* mutation testing and treatment have access to these interventions? (see prompts for *EGFR*)

6. Are you aware of any strategies to increase the number of patients receiving *EGFR* mutation targeting treatments?

7. Are there any strategies which you believe could be implemented to further increase the number of patients receiving *EGFR* targeting treatments in the future?

**Other key stratified medicines**

(Repeat above questions for other commonly used stratified medicines identified by respondent)

**Stratified medicines in trials**
Do you believe there may be factors which will slow the implementation of these newer examples of stratified medicines?

To what extent do you believe previous experience will facilitate the implementation of future stratified medicines? (e.g. individual, regional, national strategies)

**Finishing questions**

Are there any other comments you would like to make about any difficulties you have had in offering stratified medicines for NSCLC?

Do you have any other questions about this study?

Thank you. I will now stop the recording

Can I now ask you some questions about yourself?

In which NHS hospital and trust do you work?

What is your job title? **(prompt for grade if not stated)**

How many years of experience do you have in this role?
Appendix 4.5: Interview Schedule for Service Commissioners

Draft Interview Schedule – Service Commissioners

(Phrases in brackets and bold represent prompts for the interviewer)

Thank you for agreeing to be interviewed.

I am a PhD student based in MCHE at The University of Manchester. I am funded by the Wellcome Trust

I want to ask you some questions to explore your experience of the commissioning arrangements of molecular diagnostic tests and particularly those used to select specific medicines for patients with NSCLC. I would like to focus on current and historical issues.

The interview should take no longer than one hour and will be recorded. Firstly could you confirm that you have received and returned the consent form for this interview? Do you have any questions before we start?

Current

Can you tell me some examples of molecular diagnostic tests you are aware of?

In general, can you tell me about how new molecular diagnostic tests to target oncology medicines are commissioned in the NHS?

Are molecular tests commissioned at the national, regional or local level?

Can you talk me through each of these commissioning arrangements?

Are the commissioning processes for such molecular diagnostic tests different for NSCLC?

Are you aware of any current challenges with the way mutation tests are commissioned in the NHS?

In general, do you think clinicians are aware of the way such tests are commissioned?

What evidence do you think is vital to inform how molecular diagnostic tests are commissioned?

Historical

Have molecular diagnostic tests always been commissioned in this way?

(If applicable) Were there any previous issues with the way molecular diagnostic tests were commissioned?

Was there a particular example of a molecular diagnostic test that you are aware of was particular problematic in terms of the commissioning process?

Was there a particular example of a molecular diagnostic test that you are aware of was an exemplar of good practice in terms of the commissioning process?

Finishing questions

Are there any other comments you would like to make about the commissioning of molecular diagnostic test?
Thank you. I will now stop the recording

Can I now ask you some questions about yourself?

What is your job title? *(prompt for grade if not stated)*

How many years of experience do you have in this role?