# Integrating early economic evaluation into the development of Target Product Profiles for new diagnostic tests

**Paola Cocco**

Submitted in accordance with the requirements for the degree of

Doctor of Philosophy

The University of Leeds

Faculty of Medicine and Health

August 2022

The candidate confirms that the work submitted is her own, except where work which has formed part of jointly authored publications has been included. The contribution of the candidate and the other authors to this work has been explicitly indicated below. The candidate confirms that appropriate credit has been given within the thesis where reference has been made to the work of others.

*Details of the candidate's contributions to publications relating to this thesis:*

1. **Cocco P, Ayaz-Shah A, Messenger MP, West RM, Shinkins B. Target Product Profiles for medical tests: a systematic review of current methods. *BMC Medicine 2020*. doi.org/10.1186/s12916-020-01582-1**

The candidate (Cocco P) was responsible for planning and designing the study, conducting the search strategies and data extraction, as well as leading the writing of the manuscript (1). The co-authors (Shinkins B, Messenger MP, West RM) provided supervisory feedback throughout the study. Shinkins B undertook screening of 10% of abstracts identified from the search strategy. In addition, Ayaz-Shah A conducted data extraction as an independent third reviewer. Research presented in this manuscript is included in **Chapter 2** of this thesis.

2. **Cocco P, Messenger MP, Smith AF, West R, Shinkins B. Integrating Early Economic Evaluation into Target Product Profile Development for Medical Tests: Advantages and Potential Applications. *Int J Technol Assess Health Care 2021*. doi.org/10.1017/S0266462321000374**

The candidate generated the ideas for the manuscript and led the writing of the manuscript (2). The co-authors helped to formulate the ideas presented in the manuscript and provided comments during the drafting phase. Research presented in this manuscript is included in **Chapter 2** of this thesis.

**Acknowledgements**

I want to thank and express my gratitude to my supervisors - Dr Bethany Shinkins, Dr Alison Smith, Dr Kerrie Davies, Dr Michael Messenger and Professor Robert West - for their constant support and guidance provided in the last four years. Robert – thank you for always being available whenever I needed your advice. Mike – thank you so much for your contagious enthusiasm towards this project. Kerrie – a massive thank you for making the intricate world of *C. difficile* infection more accessible to me, and most importantly, more fun. Your feedback and explanations have always been invaluable. Alison – I am extremely grateful for your constant support, encouragement and your thorough and insightful feedback. A big thank you for walking me through DES modelling, and for always being available for additional clarifications. Beth – thank you so much for always giving me superb guidance, as well as enough confidence and room for growing as a researcher, and for the plenty of opportunities outside of this thesis. Thank you for always supporting me and for believing in me, especially when I did not. To each of my supervisors – Beth, Alison, Kerrie, Mike and Robert – thank you for every chat, meeting and for your time in general, especially during the COVID-19 pandemic. I have always felt you had my back.

I also would like to thank The Biotechnology and Biological Sciences Research Council, the Engineering and Physical Sciences Research Council, and the Medical Research Council for funding this research.

I would like to express my gratitude towards Professor Mark Wilcox, Dr Jonathan Sandoe, Professor Alastair Hay, Anam Ayaz-Shah, Dr Stephen Bradley, Bryony Dawkins, William Goodman, Kelly Lloyd, Maisie Martland and John O'Dwyer for helping with pre-testing and validating the survey. I am also grateful to Dr John Coia, Dr Chris Settle, Dr Helen Forrest, Dr David Jenkins and Dr Sarah Mumford for piloting the survey. Many thanks to Dr Beverly Riley and Dr Chris McKee for helping with the dissemination of the survey. I also would like to express my gratitude to Dr Christopher Rooney for helping me to understand the care pathways at the Leeds Teaching NHS Hospitals, and to Dr Armando Vargas-Palacios for the useful insights on the modelling chapters. A special thanks to Dr Gurdeep Sagoo for his endless generosity.

On a personal level, this thesis has been written while moving between two countries, the UK and Italy, because of the COVID-19 pandemic – so there are many people to thank who have made the completion of this thesis possible.

First and foremost, I want to thank my parents who have always supported me throughout my education – although this time I feel you definitely took the extra mile for me. I want to dedicate this thesis to you both, Mamma and Papà, because

without your continuous support, endless patience and care, this thesis would have been (very) far from completion. You have welcomed me back home in a moment of need, and kept me going despite everything. Words cannot express my gratitude to you both. A big thanks to my sister Laura and my relatives for cheering me up and for always listening to me. Spending time together back in Cagliari was a gift.

I also want to thank my hometown friends, Benni, Jack, Chiara, Giulia and Noce for their support, and for making the two years in Cagliari more fun and enjoyable. Many thanks to Evrim, Francis, Shaista and Kelly and the other friends from the PhD office and in Leeds for being an amazing support system and companions during this PhD.

A massive thank you to my dragon boat team in Shipley, Yorkshire Sharks, for making my time since I got back to the UK extremely fun, with plenty of activities with wonderful people. To conclude, a special thanks to Sophie, whose words of encouragement, patience and funny jokes made me go through the end of this PhD journey with a constant smile on my face. And now, let’s move on to the next chapter.

**Funding**

This thesis was written as part of a full-time School of Medicine PhD Scholarship awarded to Paola Cocco at the University of Leeds. This work is also supported by the 'Antimicrobial Resistance Cross Council Initiative' (Grant number MR/N029976/1), Funding Partners: The Biotechnology and Biological Sciences Research Council, the Engineering and Physical Sciences Research Council, and the Medical Research Council.

# Abstract

**Background**: Target Product Profiles (TPPs) for new medical tests specify desired properties and performance characteristics required to address an unmet clinical need. Early economic evaluation (EEE) has been used to identify the acceptable ranges for test specifications in the context of health technology assessment, but has not yet been utilised as part of the TPP methodology.

**Aim**: To explore how EEE methods can be integrated into the TPP development process for new tests.

**Methods**: A systematic review of TPPs for medical tests was conducted to establish current TPP methodology and limitations. The potential utility of integrating EEE into TPPs was explored via a case study example: a new rapid diagnostic test for Clostridioides difficile infection (CDI). An online survey of UK healthcare professionals was first conducted to map current practice and identify unmet clinical needs. An EEE, consisting of a discrete event simulation model, was then developed to determine minimum performance specifications for key test properties (diagnostic sensitivity, specificity) based on cost-effectiveness considerations.

**Results**: The systematic review identified a typical three-step development process for TPPs for tests – scoping, drafting and consensus-building phases. The use of subjective data sources, poor methodological transparency and an oversight of clinical utility and cost-effectiveness considerations were identified as key limitations of current TPP methodology. A three-phase *de novo* approach was developed to derive minimum test performance specifications and maximum costs using a series of sensitivity analyses. In the case of a new rapid test for CDI, a minimum diagnostic sensitivity and specificity of 96% is required for the test for CDI to be cost-effective.

**Conclusions:** EEE, in conjunction with stakeholder consultation activities, provides an evidence-based and transparent approach to informing certain TPP test characteristics. EEE should also help TPP developers to better understand how trade-offs between different test characteristics play out in terms of cost-effectiveness.

# Table of Contents

# List of Tables

# List of Figures

# Abbreviations

| | |
|---|---|
| ABS | Agent-based simulation |
| ACCE | Analytic validity, Clinical validity, Clinical utility, Ethical legal |
| AIC | Akaike Information Criterion |
| AMR | Antimicrobial resistance |
| ARR | Aldosterone-to-renin |
| BIC | Bayesian Information Criterion |
| BNF | British National Formulary |
| C. difficile | Clostridioides difficile |
| CA | Clinical assumption |
| CBA | Cost-benefit analysis |
| CCG | Clinical Commissioning Group |
| CCNA | Cell cytotoxicity neutralisation |
| CDC | Centers for Disease Control and Prevention |
| CDI | Clostridioides difficile infection |
| CDF | Cumulative Distribution Function |
| CDRN | Clostridium difficile ribotyping network |
| CE | Conformité Européenne |
| CEA | Cost-effectiveness analysis |
| CHEERS | Consolidated Health Economic Evaluation Reporting Standards |
| CHERRIES | Checklist for Reporting Results of Internet E-Surveys |
| CI | Confidence interval |
| COMBACTE-CDI | Combatting Bacterial Resistance in Europe |
| COPD | Chronic obstructive pulmonary disease |
| CRC | Colorectal cancer |
| CRF | Case report form |
| CRUK | Cancer Research UK |
| CSLI | Clinical & Laboratory Standards Institute |
| CTC | Computed tomographic colonography |
| CUA | Cost-utility analysis |
| DALY | Disability Adjusted Life Year |
| DAP | Diagnostic Assessment Program |

| | |
|---|---|
| DCE | Discrete Choice Experiment |
| DCRSHP | Diagnostic classifier for risk stratification of haematuria patients |
| DES | Discrete event simulation |
| DH | UK Department of Health and Social Care |
| DNDi | Drugs for Neglected Diseases initiative |
| ED&D | Early cancer detection and diagnosis |
| EEA | European Economic Area |
| EEE | Early economic evaluation |
| EFLM | European Federation of Clinical Chemisty and Laboratory Medicine |
| EIA | Enzyme immunoassay |
| EMA | European Medicines Agency |
| EQ-5D | European Quality of Life-5 Dimensions |
| EU | European Union |
| FDA | Food and Drug Administration |
| FIND | Foundation for Innovative New Diagnostics |
| FIT | Faecal immunichemical test |
| FN | False-negative |
| FP | False-positive |
| GDH | Glutamate dehydrogenase |
| GI | Gastrointestinal |
| GP | General Practitioner |
| GPP | Gastrointestinal pathogen panel |
| HIV | Human immunodeficiency virus |
| HRA | Health Research Authority |
| HT | Hypothetical test |
| HTA | Health Technology Assessment |
| ICER | Incremental cost-effectiveness ratio |
| ICH | International Council for Harmonisation |
| IDC | International Diagnostic Center |
| INMB | incremental net monetary benefit |
| IPD | Individual patient data |
| IQR | Interquartile range |
| IVDD | In-Vitro Diagnostic Devices Directives |

| | |
|---|---|
| IVDR | In-Vitro Diagnostic Regulation |
| LIMS | Laboratory Information Management System |
| LTHT | Leeds Teaching Hospital NHS Trust |
| LuCID | Longitudinal European Clostridium difficile Infection Diagnosis Surveillance Study |
| MACE | Major adverse cardiac event |
| MHRA | Medicines and Healthcare products Regulatory Agency |
| MLE | Maximum likelihood estimation |
| MPS | Minimum performance specification |
| MRC | Medical Research Council |
| MTG | Medical technologies guidance |
| NAAT | Nucleic acid amplifications test |
| NDPP | NHS England Diabetes Prevention Programme |
| NHS | National Health System |
| NICE | National Institute for Health and Care Excellence |
| NIHR | National Institute for Health and Care Research |
| ODE | Ordinary differential equations |
| OGPPH | Office of Genomics and Precision Public Health |
| PA | Primary Aldosteronism |
| PCR | Polymerase Chain Reaction |
| PHE | Public Health England |
| POCT | Point-of-care test |
| POCTRN | Point-of-Care Technologies Research Network |
| POSA | Probabilistic One-way Sensitivity Analysis |
| PPIE | Patient and Public Involvement & Engagement |
| PRISMA | Preferred Reporting Items for Systematic Reviews and Meta-analysis |
| PSA | Probabilistic sensitivity analysis |
| PSA test | Prostate-specific antigen test |
| QALD | Quality-adjusted day |
| QALY | Quality-adjusted life year |
| QTPP | Quality Target Product Profile |
| R&D | Research and Development |
| $R_0$ | Reproductive ratio |

| | |
|---|---|
| RCT | Randomised controlled trial |
| RID-AMR@Leeds | Rapid infections diagnostics to combat antimicrobial resistance research group |
| SD | Standard deviation |
| SD models | System dynamic model |
| SIGHT | Suspect Isolate Gloves Hand washing Test |
| SimCRC | Simulation Model of Colorectal Cancer |
| Simoa | Single-molecule array |
| SoMREC | School of Medicine Research Ethics Committee |
| STRESS | Strengthening The Reporting of Reporting of Empirical Simulation Studies guidelines |
| TAP | Technology Appraisal Programme |
| TB | Tuberculosis |
| TC | Toxigenic culture |
| TcdA | Toxin A |
| TcdB | Toxin B |
| TN | True-negative |
| TP | True-positive |
| TPP | Target Product Profiles |
| UK | United Kingdom |
| UKCA | UK Conformity Assessment |
| USA | United States of America |
| VOI | Value of Information |
| WCC | White blood cell count |
| WG-TE | Test Evaluation Working Group |
| WHO | World Health Organisation |
| WTP | Willingness-to-pay |

# Chapter 1
# Introduction

## 1.1 Chapter outline

This chapter introduces the main topics discussed in this thesis. An overview of roadmap to adoption of new medical tests into clinical practice is first presented (see section 1.2), followed by an introduction to the concept of *Target Product Profiles* (TPPs) – defined as strategic documents listing the minimal and desired properties for a given product or intervention (section 1.3). Section 1.4 introduces the key concepts underpinning economic evaluations, with a particular focus on *early economic evaluations* (EEEs) of medical tests and the role they play in supporting the development of new tests. The rationale of the thesis is then presented in section 1.5 – in particular, highlighting the potential advantages of integrating EEE into the TPP development process for medical tests. The final section presents the thesis case study, scope, aim, and structure (section 1.6).

## 1.2 Roadmap to adoption of new medical tests into clinical practice

The global trend towards an earlier and more personalised approach to medicine has been accompanied by a surge of innovation in medical testing (3). The translation of scientific discoveries from 'bench to bedside', however, is often lengthy and challenging for new healthcare technologies (3, 4), particularly so for diagnostic tests. Despite extensive research and initial optimism (5, 6), many innovative diagnostic tests, ranging from cancer biomarkers[1] (8, 9) to molecular genetic tests (10), fail to enter clinical practice (6, 11). For example, it is estimated that less than 1% of novel cancer biomarkers actually reach clinical practice (12).

The roadmap to adoption for new medical tests consists of several consecutive stages, including: (1) identification of unmet clinical need (see section 1.2.1); (2) the test evaluation pathway (section 1.2.2); (3) regulation (section 1.2.3); (4) reimbursement (section 1.2.4); and (5) adoption (section 1.2.5). The following

[1] A biomarker is defined as "a biological molecule found in blood, other body fluids, or tissues that is a sign of a normal or abnormal process, or of a condition or disease" (7).

sections describe each of these stages, alongside some of the key barriers new medical tests face throughout the journey to adoption into clinical practice.

### 1.2.1 Identification of unmet clinical need

Although the research and development (R&D) process of innovative tests should ideally be driven by patients' needs, laboratory discoveries or technology advancements often play a more prominent role in stimulating innovation (i.e. 'technology-driven innovation' rather than 'needs-driven' innovation) (13-15). The unmet clinical need addressed by novel tests is often an afterthought, and early engagement with test end-users (e.g. clinicians and patients) has historically been overlooked (14). As a result, new technologies often fail to clearly address a specific problem or need within the care pathway where the test is intended (5, 6, 13, 16). This, in turn, can result in difficulties demonstrating the clinical and economic benefits associated with a test, and ultimately failure of adoption into clinical practice (6).

Recent attempts have been made to place the concept of unmet clinical need at the heart of the R&D process for new tests – such as the Test Evaluation Working Group (WG-TE) of the European Federation of Clinical Chemistry and Laboratory Medicine (EFLM) checklist to identify unmet clinical needs for new biomarkers, published in 2016 (17, 18). This 14-item checklist aims to: (i) identify the disease area and patient population where a new biomarker is required; (ii) verify if there is an existing solution to fulfil such unmet clinical need; (iii) define how (and if) the new biomarker would enhance current clinical practice; and (iv) assess the technical, economical and organisational feasibility of using the new biomarker in practice. This checklist, however, does not define how to translate an identified unmet clinical need into a test technology solution – especially in regards to the technical and performance specifications new tests should possess to address a given need. Nevertheless, two studies using the EFLM checklist to establish the case for new diagnostic tests have now been published, including tests for cardiovascular risk stratification (19), and acute kidney injury (20). It is expected that awareness and uptake of such tools will increase over time, ensuring that clinical needs are at the heart of R&D activities going forward.

### 1.2.2 Test evaluation pathway

New tests undergo a lengthy development, validation and evidence generation process, starting from the discovery of a target biomarker and culminating with the evaluation of the broader impact of introducing a new test into clinical practice (21). Several published frameworks summarising the key evidence requirements for new tests are available (22), including prominent examples such as the Analytic validity, Clinical validity, Clinical utility, Ethical legal and social implications (ACCE) framework published in 2003 by the Centers for Disease Control and Prevention (CDC)'s Office of Public Health Genomics (OPHG) (23), and the evaluation framework from the WG-TE of the EFLM published in 2014 (21). More recent frameworks have also been published (8, 24), but have largely included the same central evidence as outlined in the ACCE and EFLM WG-TE frameworks (summarised further below).

Prior to commencing test evaluation activities, it is important that the *test purpose* and *role* have been defined. *Test purpose* refers to the intended clinical application of a new test (e.g. (i) diagnosis; (ii) prognosis; (iii) monitoring; (iv) screening) (20), and how the test result will inform patient management (21). *Test role* meanwhile, refers to the mechanism by which the test is intended to alter current practice – e.g. a new test may *replace* a current standard care test; may be used as an initial *triage* test; or may function as an *add-on* test in additon to current tests in place (25). The specified role and purpose of a test influence the specific approaches that will be required throughout the test evaluation pipeline.

The key evidentiary requirements typically demanded for new tests are described in the *test evaluation pathway* as illustrated in Figure 1-1 and defined below:

1. **Scientific validity** – ensures that: (i) there is a proven association between a newly found biomarker and a certain disease; and (ii) that the new test is able to measure that biomarker. This stage of research typically entails scientific proof-of-concept studies.
2. **Analytical performance** – is the ability of a test to correctly detect and measure the analyte[2] of interest. This stage of research typically entails laboratory-based studies to evaluate key analytical performance metrics of a new test – including analytical sensitivity and specificity, measurement

[2] An analyte is defined as "a material or substance the presence or concentration of which in a specimen is determined by analysis" (26).

bias and precision, and limits of detection – against a reference measurement method.

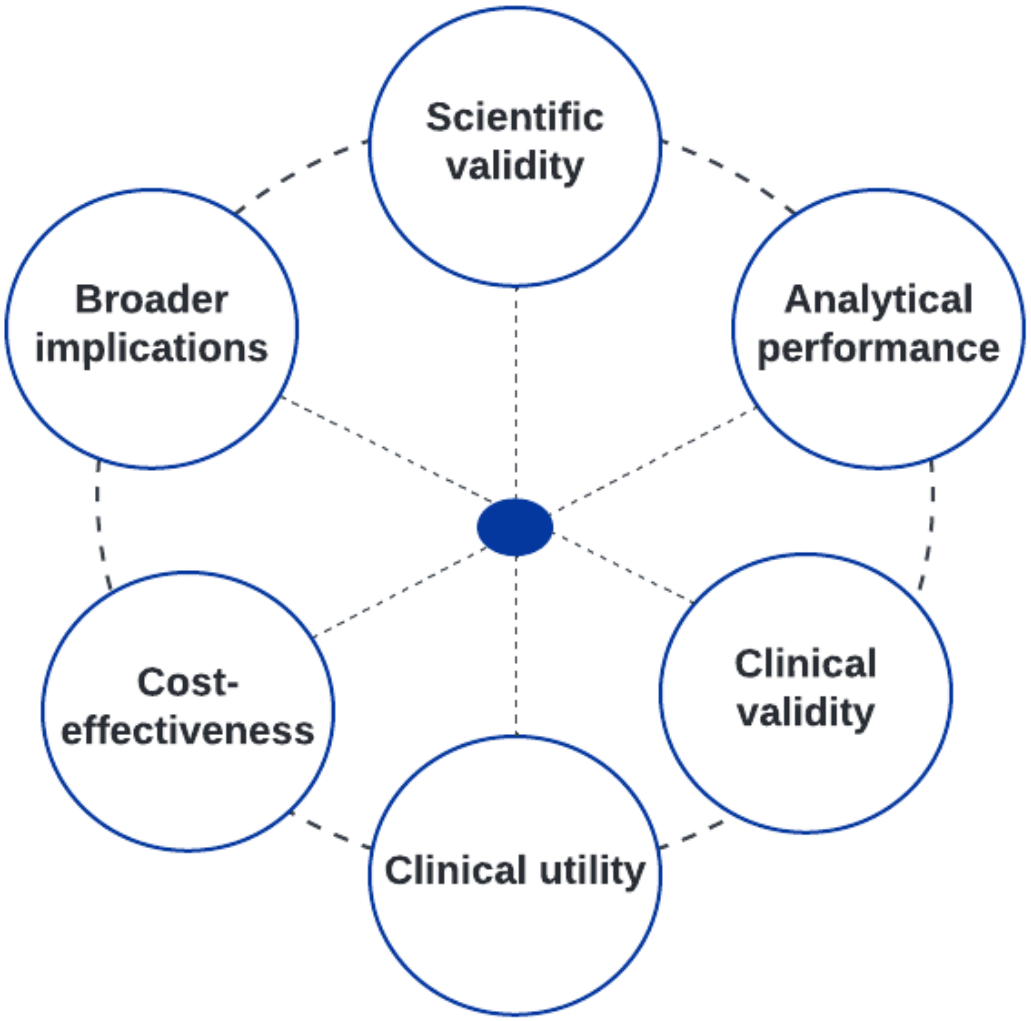
3. **Clinical validity** – is the "ability of a test to yield results that are correlated with a particular clinical condition of a physiological or pathological process or state" (27). In relation to diagnostic and screening tests, evaluation of the clinical validity focuses on assessment of diagnostic accuracy – which is defined as the ability of a test to differentiate between those with and without the target condition (as measured by diagnostic sensitivity and specificity)(28). Ideally diagnostic accuracy should be prospectively evaluated based on samples taken from a representative cohort of patients and compared against a diagnostic reference method.
4. **Clinical utility** – concerns the clinical impact of introducing a new test on: (i) patient management decisions (e.g. decision to treat or not to treat) and (ii) patient health outcomes (e.g. improvement in quality of life, longer survival, fewer adverse events). A particular issue that hinders evaluation of medical tests concers the difficulty in establishing clinical utility. Due to the indirect nature in which test results impact on patients' health (i.e. via clinical management and treatment decisions, rather than directly impact on patient health outcomes) (21), measuring and proving the clinical utility requires evidence of both: (i) how the information from a test is incorporated into decision-making; and (ii) the downstream effectiveness of those decisions (29). This information might not be available in the context of a novel testing technology. Evaluation of clinical utility is ideally evaluated via a high quality randomised controlled trial (RCT). This, however, is often not feasible for a number of reasons (e.g. length of follow-up required, sample size requirements, not ethical to randomise, too many testing comparators to include) and alternative methodologies are required – such as decision-analytic modelling (30).
5. **Cost-effectiveness** – relates to the comparative analysis of the costs and health effects associated with the new test compared to current standard practice. It informs whether the new testing strategy can be regarded as good 'value for money'. For further detail on the principles underpinning cost-effectiveness analysis and economic evaluation, see section 1.4. Common approaches to economic evaluation include trial-based assessments (in which cost and health outcomes impacts are estimated

based on clinical trial data) and model-based assessments (in which cost and health outcomes are estimated using a decision-analytic model synthesing data across multiple sources).

6. **Broader implications** – concern the broader impact and consequences of introducing a new test technology into clinical practice – in terms of expected acceptability, usability, psycological, legal, ethical, societal, and/or organisational consequences. The broader impact of introducing new diagnostic tests is often overlooked during test evaluation (31), as well as human factors and ergonomics (32). Aspects related to human factors (e.g. acceptance, usability, user experience and expectations to the test) are usually measured with standardised questionnaires, direct observations, interviews, standardised scales and post-use data gathering methods (e.g. diary) (32). Current study designs and methods, however, lack an integrated approach which captures simultaneously different aspects related to the use of a new test – such as expectations, user experience and the context in which the test is used (32).

Evidence generation for tests can take many years. A systematic overview of diagnostic horizon scan reports found all evidence requirements for new point-of-care tests (POCTs) had been collected on average over 9 years (IQR 5.5-12.5 years) (31) – with triage tests requiring 15 years (IQR 10-19) to collect all evidence components, as opposed to 9 years (IQR 5-11) for replacement tests. A key concern is that, as technology and innovation is developing at a rapid pace, novel technologies may be obsolete by the time they have been fully evaluated for use in clinical practice.

**Figure 1-1 Key components of the test evaluation pathway. Figure adapted from Horvath, A. et al. (21)**




### 1.2.3 Regulation

Before being introduced into the market, new tests must receive regulatory approval to ensure compliance with regulatory standards in place across where manufacturers wish to launch their new test.

In the European Union (EU) and the UK, manufacturers must currently obtain the Conformité Européenne (CE) mark to introduce new tests into the European Economic Area (EEA) or UK market. The CE mark indicates that the new test complies with the European In-Vitro Diagnostic Devices Directives (IVDD 98/79/EC) (33). In 2017, the new European In-Vitro Diagnostic Regulation (IVDR 2017/746) (27) was proposed, coming into full force in 2027 (34), replacing the IVDD.

Core to the IVDR legislation is to maximise the real-world effectiveness of new tests being adopted into the market, as well as prioritising patient safety, throughout the entire lifecycle of the test (35). Compared to the IVDD, the newer IVDR legislation requires manufacturers to fulfil more stringent requirements in relation to the technical documentation and post-market surveillance of tests (35). Manufacturers are required to provide detailed evidence that their new tests meet pre-determined performance characteristics relating to: (i) scientific validity, (ii) analytical performance (e.g. analytical sensitivity and specificity, trueness, precision); and (iii) clinical performance (e.g. diagnostic sensitivity and specificity, predictive values, likelihood ratios) (27).

It should be noted that the EU IVDR legislation will not come into effect in the UK, although CE marking will be recognised in the UK until 30th June 2023. Subsequently a separate UK Conformity Assessment (UKCA) will be required for new tests to be marketed into the UK. Current details as to how the UK legislation will deviate from the EU IVDR legislation is as yet unknown.

### 1.2.4 Reimbursement

After undergoing validation and regulatory processes (as described above), the adoption of a new test into routine clinical practice is ultimately dependent on local and national reimbursement authorities. In the context of increasing healthcare expenditures and constrained budgets, local and national reimbursement authorities face challenging priority-setting and resource allocation decisions (6). Reimbursement processes aim to ensure that the proposed new test is expected to be of value to the patient population of interest and the health service (36). As part of the reimbursement process, the clinical and cost-effectiveness of a new test against the standard testing strategy is assessed to ensure that the new test offers good ‘value for money’ (6).

Test adoption and reimbursement decisions are informed by health technology assessment (HTA), defined as

> “the systematic evaluation of properties, effects and/or impacts of a health technology. It is a multidisciplinary process to evaluate the social, economic, organizational and ethical issues of a health intervention or health technology. The main purpose of conducting an assessment is to inform policy decision-making” (37).

In England, multiple routes to reimbursement are available for new medical tests, including

- **receiving national guidance (but not mandated reimbursement) from the National Institute for Health and Care Excellence (NICE)** – new medical tests may be evaluated under three different NICE schemes, depending on the expected role and impact of the test: (i) the Technology Appraisal Programme (TAP); (ii) the diagnostic assessment program (DAP); and the medical technologies guidance (MTG) (38). Developers are usually required to submit cost-effectiveness evidence of their new testing technology under current NICE evidentiary requirements.
- **approaching local health decision-makers** – such as National Health System (NHS) hospital trusts or Clinical Commissioning Groups (CCGs). While the evidentiary requirements demanded by local decision-makers appears to be relatively unknown (39, 40), a qualitative study found that funding decisions for diagnostics are based on several factors, including 'value for money' considerations.

Whilst there are multiple routes to reimbursement, each requiring different specific evidence, evidence on all of the components of the test evaluation pathway (e.g. cost-effectiveness) is typically sought by all of these decision-makers.

### 1.2.5 Adoption

Following a positive reimbursement decision, the adoption of a novel testing technology into day-to-day clinical practice ultimately depends on the clinicians' acceptance and their widespread uptake of the new technology and integration into the healthcare systems (41).

When offered new tests with proven superior value, some clinicians may nevertheless show reticence in embracing the new technologies (41-43). For example, clinicians may be wary of new technologies they have no personal experience with, and may be unwilling or unable to undergo additional training required to confidently utilise the new test and interpret its results. (41).

Alternatively, there may be a reluctance among some clinicians to change the status quo of clinical practice, thereby making it difficult to implement innovations and de-adopt existing technologies (44). A paucity of data on the excepted

benefits and harms underpinning a new testing option might also explain a resistance among some clinicians towards implementing innovative testing technologies (44).

One potential way of improving clinician buy-in is to ensure that stakeholder preferences (including clinicians and patients) are considered throughout the whole R&D process of a new test. Potential barriers to adoption can thus be identified in earlier stages of the evaluation pathway, and processes put in place to deal with them. This would increase the likelihood of new tests being adopted and implemented into clinical practice (6).

## 1.3 Target Product Profiles

Based on the discussed challenges in introducing new tests into clinical practice, there is a clear need for new methods and processes to ensure that innovation and research efforts are focused on technologies that fulfil a specific clinical need, and have the potential to bring value to patients and the health care system.

To this end, Target Product Profiles (TPPs) have been suggested as a means of accelerating and de-risking the innovation process for new diagnostics, and ensuring that tests fulfil an unmet clinical need (45). TPPs are strategic documents listing the desired properties for a given product or intervention (46). While a universal definition for TPPs and understanding of their use are currently lacking (46, 47), there are core common features across TPPs to date, including

- **the concept of 'beginning with the goal in mind'** – TPPs establishes key features and performance specifications in advance of product development, to ensure that the new technology is developed in a way that meets specific goals and expectations (48, 49); and
- **the concept of 'living' documents** – TPPs should ideally be refined and updated as additional relevant information becomes available, to ensure that performance specifications listed in TPPs reflect the current evidence base (46, 48).

Historically, there have been three applications of TPPs, including: (1) in-house industry-led documents; (2) non-mandatory regulatory documents for pharmaceuticals; and (3) guiding documents in the public health sector (46, 50). The following sections describe each of these applications.

### 1.3.1 Industry-led application of TPPs

In the 1950s and 1960s, TPPs were initially developed in-house by industry, specifically for chemical and electronic products (50). With the widespread adoption of project management practices in the 1980s and 1990s, the use of industry-led TPPs began to be increasingly recognised as a means of identifying in advance project goals and potential risks (49, 50). In this context, TPPs are used either as:

- **communication tools** – TPPs facilitates communication of how a certain product is used, alongside an overview of the key properties of the product, between different groups within the company (e.g. R&D department, manufacturing, marketing) throughout the product development process (49, 50); or
- **strategic documents** – TPPs can also be used as a basis for comparing competing products against the product being developed. This helps to inform pricing strategies (46), and to understand how a product might differ from potential competitors already in the market (47).

### 1.3.2 Regulatory applications of TPPs

In the USA, TPPs are used as voluntary briefing documents to stimulate discussion between the pharmaceutical manufacturers and the USA Food and Drug Administration (FDA) throughout the drug development process (48). In this context, TPPs contain key goals for the manufacturer to achieve (48), and help to fulfill two regulatory objectives (51). First, TPPs support a discussion thoroughout the drug development process about the expected clinical and technical requirements the new treatment should achieve. Within a TPP document, manufacturers can also list the studies intended to address the *labelling concepts*, be it either planned or concluded studies (48). Second, TPPs help to gain feedback and stimulate discussion on the attributes of a certain drug once finalised. In the USA, guidance for developing TPPs is available in relation to new pharmaceutical drugs (48). This guidance, issued by the USA FDA, provides an overview of the purpose and attributes of TPPs, and which requirements for a new drug should be included based on safety and efficacy considerations (48).

Although manufacturers are not legally required to submitt TPPs, anecdotal evidence indicates that TPPs can be used to support discussion with manufacturers, helping to minimise late-stage product development failures, whilst ensuring a rapid generation of safety and efficacy data to meet the drug development program's requirements (48, 50). Published literature, however, suggests that TPPs are seldom submitted for discussion with the regulatory body and, when submitted, they are usually incorporated at very late stages of the development process (50, 51).

In addition to this, the International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use (A)[3] guideline Q8 refers to TPPs as a means of supporting effective communications between regulatory bodies and industry throughout the drug development process (53). Core to this guidance document is the concept of *quality by design*, which is defined as:

> "A systematic approach to development that begins with predefined objectives and emphasizes product and process understanding and process control, based on sound science and quality risk management" (53).

Under the *quality by design* framework, a new product is designed with the aim of meeting pre-identified quality objectives (54, 55). One of the first steps underpinning the *quality by design* framework is to develop a TPP, also known in this context as a *Quality Target Product Profile* (QTPP) (55). A survey study found that industry pharmaceutical representatives regard the *quality by design* principles as beneficial for the patients and the manufacturing process as a whole, although a subset of respondents reported that they have never used QTPPs (56).

### 1.3.3 Guiding documents in the public health sector

In addition to being industry-led and regulatory documents, TPPs have also been used as guiding documents within public health organisations as means of transparently identifying the necessary features of a public health intervention

[3] In 1990, the ICH was created in Brussels with a mission "to achieve greater harmonisation worldwide to ensure that safe, effective, and high quality medicines are developed and registered in the most resource-efficient manner" (52). ICH therefore develops guidelines in conjunction with scientific and regulatory expert input.

(46). In this context, the purpose of a TPP is to identify, upfront, the minimal and desirable features and performance specifications for new technologies to fulfil a pre-specified unmet clinical need (57, 58). This helps to coordinate and optimise the efforts and resources of every stakeholder and actor involved in the development process to achieve an intended health-related goal (46, 58). An example of TPPs in this context is the PATH Malaria Vaccine Initiative, a multidisciplinary project which aimed to develop TPPs for candidate malaria vaccines (59). The application of TPPs as guiding documents in the context of medical tests is discussed in the section below.

### 1.3.4 TPPs for medical tests

Outside of the development process of pharmaceutical drugs and vaccines, TPPs can also be useful for designing 'fit for purpose' medical tests (60). In this context, TPPs usually report the clinical purpose of the test, the target population and healthcare setting, in addition to minimal and optimal performance benchmarks, technical and infrastructural requirements (e.g. power requirements, storage conditions and waste disposal) (60). Required test characteristics may be expressed either quantitatively (e.g. limit of detection, diagnostic accuracy) or qualitatively (e.g. intended test use, target user). Table 1-1 illustrates a hypothetical example of TPP for a rapid diagnostic test for tuberculosis (TB). For each test characteristic (e.g. target user, price of individual test, analytical sensitivity), acceptable and desirable values are reported, and the sources used to derive those values are listed in the reference column (e.g. literature, consultations with experts).

TPPs have great potential to be used as guiding documents for tests developers to avoid late-stage development failures and reduce research waste. They can be used during the development and manufacturing phase to ensure that a new test meets pre-established operational and performance requirements, in line with a specified unmet clinical need (48). This ensures that innovation of new medical tests is driven by patient and health service's needs, rather than solely technological discoveries. This should increase the likelihood that new tests will be adopted into clinical practice and reimbursed (6, 31). Nevertheless, despite the relevance of TPPs for test manufacturers, there is currently no formal guidance as to best practice methods for developing a TPP specifically for medical tests and regulatory applications in this setting are, so far, limited.

**Table 1-1 Hypothetical example of a TPP for a rapid test for TB**

| Characteristic | Acceptable (*'must have'*) | Desired (*'would like to have'*) | Reference |
|---|---|---|---|
| **Intended use** | Rapid diagnosis of TB to initiate early TB treatment | | Expert consensus |
| **Target population** | Adults suspected of having active TB | | Expert consensus |
| **Target healthcare setting** | Secondary care with on-site laboratories | Primary care clinics without access to on-site laboratories | Expert consensus |
| **Target user** | Healthcare worker | Trained lay person | Expert consensus |
| **Diagnostic sensitivity** | ≥ 85% | ≥ 95% | Literature |
| **Diagnostic specificity** | ≥ 80% | ≥ 90% | Internal/unpublished data |
| **Unit test price** | £40 | ≤£10 | Literature |

## 1.4 Economic evaluation

### 1.4.1 Main concepts

Economic evaluation, a core component of HTA and the test evaluation pathway, compares the costs and benefits of two or more alternative healthcare technologies or pathways to support adoption and reimbursement decisions (61). In the context of HTAs, there are two primary types of economic evaluation used to inform test adoption decisions, which differ in terms of the measure of health benefit evaluated:

- **Cost-utility analysis (CUA)** – uses a generic measure of health gain, called the Quality Adjusted Life Year (QALY) (61, 62). A QALY is a composite measure of health which combines both quality and quantity of life (61);
- **Cost-effectiveness analysis (CEA)** – evaluates health gains in natural units (e.g. life-years gained, infections avoided) (61). CUA may be considered a subset of CEA.

In addition, depending on the perspective of the analysis and decision-maker, different sets of costs can be included – be it either healthcare (i.e. treatment and medical costs for managing the disease) or societal perspective (i.e. transportation costs and time off work are also included) (61).

Economic evaluation identifies which healthcare interventions and/or technologies should be recommended for a certain patient population based on cost-effectiveness considerations (61). Assessing whether a certain intervention represents ‘value for money’ over a comparator depends on a specific cost-effectiveness threshold, which represents the maximum willingness-to-pay (WTP) for an additional unit of health benefit (63). Different WTP thresholds may be required depending on which metric of health benefit is used. In the UK, NICE guidelines recommend a WTP threshold ranging from £20,000 to £30,000 per QALY gained (62, 64).

Two primary metrics can be evaluated when estimating the cost-effectiveness of an intervention (A) compared to a selected comparator (B), including

- **Incremental cost-effectiveness ratio (ICER)** – this ratio reflects the difference in costs (ΔC) relative to the difference in effectiveness (ΔE), either expressed as QALYs or in natural units, for A vs B (61, 63). The ICER between intervention A and intervention B is calculated as follows:

$$ICER = \frac{C_A - C_B}{E_A - E_B} = \frac{\Delta C}{\Delta E}$$

- **Incremental net monetary benefit (INMB)** – converts the incremental benefits of an intervention into costs, using a specified WTP threshold. The incremental costs are then substracted to provide the net value of monetary benefit (65). The INMB between intervention A and intervention B is calculated as follows

$$INMB = (\Delta E \times WTP\ threshold) - \Delta C$$

Depending on the outcome measure selected, an intervention is considered cost-effective compared to the comparator if: (i) the ICER falls below a specific WTP threshold (61); or (ii) if the INMB is greater than 0.

A standard approach used in economic evaluations – early or otherwise – is *decision-analytic modelling*: a framework of analysis which uses mathematical tools and schematics to provide a simplified representation of the decision problem, as a series of uncertain events. Decision-analytic modelling is often used to evaluate diagnostics in the absence of RCT data or when direct evidence on downstream benefits or harms is lacking (66). Modelling also facilitates the exploration of uncertainty surrounding test characteristics, which can provide key insights at the early stages of evaluation when evidence may be missing or highly uncertain (29). Different modelling techniques are available to conduct economic evaluations, including decision trees, Markov models, and discrete event simulation (DES) models. An overview of the available decision modelling techniques is included in Appendix A.

### 1.4.2 Early economic evaluation

Whilst economic evaluation is usually conducted at later stages of test evaluation, or after product development (67), EEE can also be performed when tests are still early development stages (e.g. from proof-of-concept stage up to the preclinical stage), when there is little or no clinical evidence available (68, 69).

EEE has been recommended as a means of guiding the development of new medical tests and for checking that a new test has the potential to be cost-effective prior to and during the typically lengthy development, regulatory and evidence generation pipeline (29). Although an universal definition for EEE is currently lacking (70) and there is variability in the methods adopted (71), a general framework for early-stage economic models for medical tests has been published (72) – see Figure 1-2. This framework comprises of five steps, including:

- **narrowing down the scope** – this initial phase aims to define the scope of the analysis in terms of: (1) the intervention(s) being evaluated; (2) potential applications of the intervention (e.g. test role and purpose); (3) the patient population intended to be tested; (4) relevant comparators to include; (5) the current and proposed clinical care pathway within which the new test(s) will function; and (6) outcomes to assess. In case the application of the new test is unclear and/or uncertain, discussions with stakeholders can be undertaken to help to clarify how and under what conditions a test could provide greatest benefits to the population of interest.
- **inventory of available evidence and data on current testing strategy** – this phase focuses on gathering evidence on the current standard care testing strategy, the clinical care pathway and any existing decision models which may help to inform the analysis. Consultations with experts are helpful at this stage to gain an understanding of the clinical pathway, and to help inform model parameters where there is a paucity of published data.
- **developing and modifying a conceptual model** – a flexible conceptual model is developed during this phase to account for different scenarios underpinning the impact of the new test – which is unknown at the outset. As opposed to late-stage economic evaluations, the conceptual structure

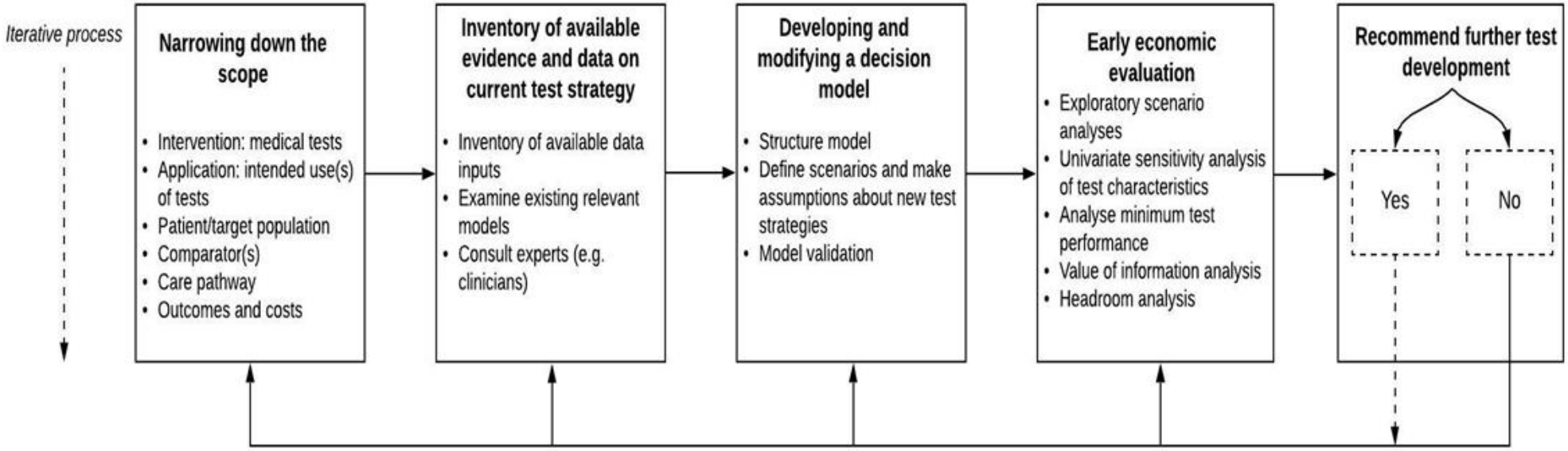
and parameters of an early-stage model are highly uncertain and could be updated and modified once new evidence emerges.

- **early cost-effectiveness analysis** – this phase aims to estimate the potential clinical- and cost-effectiveness of a new test compared to the standard care testing strategy via a series of sensitivity and scenario analyses. The findings of this phase are exploratory in nature rather than definite – as per late-stage economic evaluations – given the likely uncertainty underpinning model parameters and assumptions.
- **recommendations regarding further test development** – as opposed to late-stage economic evaluations which aim to inform reimbursement decisions, results from early economic models can be used to (68, 71, 72): (a) inform investment and design decisions of medical tests alongside future research design – if a new test is unlikely to be cost-effective at early development stages, the developers should consider suspending the development process as the likelihood of the test being reimbursed by decision-makers is low; (b) identify key evidence gaps and thereby guide future research design; (c) understand the mechanisms through which a new test has an impact on patient health outcomes; (d) establish performance requirements; and (e) plan reimbursement strategies.

A particularly useful component of early economic modelling is the process of graphically mapping the existing clinical pathway, and comparative pathways which integrate one or more novel interventions, as a series of processes (73). As the expected benefits of a new test are rooted in the clinical pathway within which the new test will sit (17, 18), this exercise helps to better articulate at early development stages the expected ways in which a new test can improve upon the current testing workflow. For example, *care pathway mapping* helps to identify the ways in which the introduction of a new test may lead to changes in clinical decisions, subsequent patient management, health outcomes, resource use and cost. This, in turn, informs the different *value propositions* for a new test – defined as the excepted monetary and non-monetary benefits for key stakeholders, including patients, clinicians and investors (73, 74). In addition, this exercise helps to understand the positioning of a new test within an existing clinical pathway (i.e *test role*) (21). These comparative pathways then underpin the structure of the decision model, where all existing evidence is synthesised to

explore the potential cost-effectiveness of the proposed technology in addition to identifying key evidence gaps.

**Figure 1-2 Steps for conducting an early economic evaluation for new tests. Figure adapted from Buisman, L.R. et al. (72).**



### 1.4.3 Previous examples of EEE for medical tests

Outside the context of TPPs, EEE has previously been used to identify acceptable ranges for test characteristics or components of the *use case*[4] (36, 75-87). Key approaches used to this end among published EEEs for medical tests are discussed below (36, 75-87).

#### Scenario Analysis

Scenario analysis can be conducted to explore the impact of a group of parameter changes, that represent a specific clinical scenario, on the cost-effectiveness of new tests. For example, scenario analyses could be used to vary the country of interest (76, 87), altering the role of the test within the care pathway (77), changing the target patient population (79), exploring different levels of compliance to test results (81), and varying the test screening age (84). Scenario analysis has also been used to test the impact of particular assumptions underlying the natural history of disease progression (80), the test-treatment pathway (75), and the inclusion of external costs of antimicrobial resistance (75).

#### Sensitivity analysis

Deterministic sensitivity analysis involves varying the values of one or more test characteristics (or other input parameters) at the time and recording the impact on the cost-effectiveness of new tests (88). For example, several authors have employed univariate sensitivity analysis (i.e. varying one parameter at a time in the model) to explore the impact of different disease prevalence values (36, 82, 85, 87), health-related utility weights (75, 85, 86), and elements of disease progression (75, 83, 86). Others have used this approach to identify ranges of minimal benchmarks of test performance such as diagnostic sensitivity and specificity (79, 82, 87), while exploring changes in cost-effectiveness results due to altering the test reference method (86), or price (75, 80, 87), alongside treatment costs (75, 87). Bivariate sensitivity analyses (i.e. varying two parameters in the model) have similarly been used to assess the impact of changing disease prevalence and test price (85), or diagnostic sensitivity and specificity (75, 77, 80, 84, 86), or diagnostic accuracy and test price simultaneously (87). In addition, several authors have conducted probabilistic

[4] *Use case* refers to the way in which medical tests are used optimally.

sensitivity analyses (PSA) in which all parameters are varied based on probabilistic distributions representing the sampling uncertainty around those parameters (71, 76, 78, 79, 81).

### Headroom Analysis

The *headroom price* of a new test represents the maximum price at which that test is considered cost-effective, assuming perfect accuracy. Headroom analysis has been used to assess the financial room for improvement compared to current practice, by either comparing the net monetary benefits between testing options (77, 78), or monetizing the effectiveness gap (82). The latter equals the maximum room for improvement between testing practice and a perfectly accurate test (82).

### Threshold Analysis

Threshold analysis identifies the critical value for input parameters above or below which a reimbursement decision is expected to change. This approach has been used to estimate the maximum price at which a test remains cost-effective given a specified WTP threshold (76, 77, 82-85). Others have used this approach to further explore the impact of changing two or more test characteristics (e.g. diagnostic sensitivity and specificity) on the test's maximum acceptable costs (76, 77, 79, 82, 83).

### Value of Information Analysis

Value of information (VOI) analysis can be used to assess the value of gathering additional information on model parameters, in order to reduce current decision uncertainty. In the context of EEE, where uncertainty around test characteristics is typically high (77, 78, 85), the utility of standard VOI approaches (e.g. computing the expected value of perfect information) may be limited, since one would always expect there to be value in conducting further research at this stage. Nevertheless, more sophisticated VOI analysis (e.g. computing the *expected value of sample information*) could be usefully employed at later development stages to make recommendations around the specific research design (e.g. sample size, duration of follow-up) expected to reduce decision uncertainty most efficiently.

## 1.5 Thesis rationale

At the early stages of development, TPPs provide test developers with key performance and technical requirements new (and existing) tests should possess to address an unmet clinical need. EEE has been used outside the context of TPPs to identify acceptable ranges for certain performance requirements for new tests (e.g. diagnostic accuracy) or components of the use case – as discussed in section 1.4.3. Selection of a WTP threshold can be used to derive the minimum performance requirements and maximum costs for a new test to be cost-effective (72). In addition, decision-analytic modelling has the potential to estimate the clinical utility of a new or hypothetical test in terms of life years, quality of life, and adverse events through linked evidence modelling. By combining EEE and WTP as a decision rule, it is possible to identify under which conditions a test maximises clinical utility (albeit with inevitable, but quantifiable, uncertainty). This illustrates the potential utility of EEE to inform TPP development. Currently, however, EEEs are yet to be used in the context of TPP development, and there is no guidance available as to how this could or should be done. Based on this, there is a clear value in exploring how early economic modelling could inform the development of TPPs for medical tests. This is expected to help manufacturers to develop tests that are both beneficial to patient's health and cost-effective, thus improving the likelihood of adoption into clinical practice.

## 1.6 Scope, aim and structure

### 1.6.1 Scope

While TPPs have been extensively produced to guide the development process of pharmaceutical drugs (as highlighted in section 1.3.2), the focus of this thesis lies in the use of TPPs in the context of any type of medical tests.

In particular, this thesis focuses on the use of EEE methods within TPPs as a means of informing key TPP performance requirements, for example diagnostic accuracy, turnaround time[5], and cost. To explore the use of EEE methods within TPP development, a case study is presented in this thesis focusing on a new rapid in-vitro point-of-care diagnostic test for Clostridioides difficile infection (CDI). Development of a full TPP for new diagnostic tests for CDI (including a wide rage of additional elements such as analytical performance and infrastructural requirements) is part of a broader Medical Research Council (MRC)-funded programme grant (MR/N029976/1)[6], and is outside the scope of this thesis. The findings from this thesis will be used at later stages to inform parts of the full TPP developed within the MRC-funded project.

As there is high variability in the incidence rates and testing protocols for diagnosing CDI across Europe (90), the scope of this thesis is on the UK clinical and diagnostic pathway for patients suspected and confirmed with CDI. Whilst the example used herein relates to infection diagnostics, and specifically an *in-vitro diagnostic tests* (i.e. tests conducted on samples derived from the human body such as blood, stool or tissue), it is expected that the general principles and methodology presented in this thesis will be applicable to most types of medical testing technologies and indications. The generalisability of the thesis findings is discussed in Chapter 7.

[5] It represents how long it takes to yield test results after having obtained the sample to test.

[6] This programme grant is called 'Rapid infections diagnostics to combat antimicrobial resistance' (RID-AMR@Leeds) research group (89).

### 1.6.2 Aim

The overarching aim of this thesis is therefore to explore how EEE methods could be integrated into the TPP development process for new medical tests. To this end, the following specific objectives will be addressed:

1. to identify and report how TPPs for medical tests are currently developed, as well any limitations with current TPP methodology;
2. to exemplify the potential benefits of using EEE to inform TPPs;
3. to identify current problems and unmet clinical needs in the diagnosis of CDI, and whether there is an unmet need for a new test for CDI; and
4. to develop an early economic model comparing a hypothetical rapid test for CDI to current practice – to identify minimum performance specifications for key parameters (i.e. diagnostic accuracy, turnaround time, test cost).

### 1.6.3 Structure

The first part of this thesis focuses on exploring how to integrate EEE methods within the TPP development process: Chapter 2 is focused on establishing current TPP methodology and key methodological limitations that EEE could help to address. The second part of the thesis from Chapter 3 to Chapter 6 applies EEE methods within the context of the case study.

The thesis is divided into seven chapters, outlined below:

- **Chapter 1** (the current chapter) provides an introduction to the development process of new diagnostic tests, TPPs for medical tests, cost-effectiveness analysis and EEE methods.
- **Chapter 2** reports a systematic literature review of TPPs for medical tests, which aims to identify the methods currently used to develop TPPs. This is followed by an overview of the anticipated benefits of integrating EEE methods within the typical TPP development process.
- **Chapter 3** introduces the clinical background to the case study. The main aspects underpinning the laboratory diagnosis of CDI are outlined, followed by a description of the UK clinical pathway for patients suspected and confirmed with CDI.

- **Chapter 4** reports an online survey of UK healthcare professionals which aims to identify unmet clinical needs and challenges in the diagnosis of CDI, and to map the clinical care pathway for CDI .
- **Chapter 5** describes the structure and parameterisation of a *de novo* early economic model for a new rapid CDI diagnostic test based on synthesising evidence from the literature, databases and consultations with clinical experts.
- **Chapter 6** presents a *de novo* pragmatic approach to identify in advance minimum performance requirements for tests based on cost-effectiveness considerations. This approach is then applied to the model-based EEE outlined in **Chapter 5**. This chapter presents the results of the model analysis, alongside a discussion of the model findings and limitations.
- **Chapter 7** summarise the key findings of the presented research and provides a discussion on the impact and limitations of the thesis, as well as future research recommendations.

# Chapter 2

# Target Product Profiles for medical tests: a systematic review of current methods

## 2.1 Chapter outline

Chapter 1 introduced the thesis topic of TPPs for medical tests as a means of stimulating 'fit for purpose' innovation for new diagnostics. Although there is a growing body of published TPPs for medical tests, no formal guidance for best practice methods is currently available. The aim of this chapter is therefore to review and describe the methods currently used to develop TPPs for medical tests, focusing on: (1) determining whether there is a common methodology used across published TPPs; (2) which test characteristics commonly featured in TPP specifications; and (3) key areas requiring further methodological development. This chapter first describes the review methods (section 2.2), followed by the results (section 2.3), and discussion (section 2.4). Based on the methodological limitations identified in the review, the hypothesis that EEE could strengthen the methodological rigour of TPPs for medical tests is presented (section 2.5).

The research and discussion presented in this chapter have been published as part of two jointly-authored peer-reviewed papers: (i) a systematic review article in *BMC Medicine* (1); and (ii) a 'Perspective' publication in the *International Journal of Technology Assessment in Health Care* (2).

## 2.2 Methods

A systematic review was conducted to explore the methods currently used to develop TPPs for medical tests. The protocol for the systematic review was registered on the PROSPERO database (CRD42018115133) (91).

### 2.2.1 Literature search

A search strategy (available in Appendix B) was developed with support from an Information Specialist (Natalie King) at the University of Leeds. The following electronic databases were searched: MEDLINE (Ovid), EMBASE, CAB Abstract Online, CINHAL, Global Health, Scopus and Web of Science. The database searches were performed in November 2018 and encompassed two main elements: (1) free-text terms to identify TPPs (e.g. quality by design, *QTPP*) (lines

#1-3); (2) MeSH headings and free-text terms related to medical tests (lines #5-6).

In addition to the electronic databases, grey literature and websites were searched using structured methods proposed by Godin, K. *et al.* (92). A customized Google search was conducted in November 2018 to find relevant websites. Each identified website was hand-searched to identify potentially relevant references or, alternatively, an internal search engine was used if available. Duplicates across searches were removed.

All searches were conducted by the primary reviewer (Cocco P) and peer reviewed by an Information Specialist (King N).

### 2.2.2 Screening

The inclusion criteria for the systematic review are presented in Table 2-1. Only references which included a TPP for a medical test were included. A restriction on the publication format was applied to ensure that the methods for each included TPP were reported in sufficient detail to review them.

**Table 2-1 Systematic review – inclusion criteria**

| Item | Inclusion criteria |
|---|---|
| **Study design** | TPP |
| **Intervention** | Any type of medical test (e.g. imaging, in vitro and in vivo medical tests) |
| **Population** | Any human population |
| **Clinical setting** | Any |
| **Indication** | Any |
| **Date** | Any |
| **Language** | English |
| **Format** | Any publication format (e.g. peer-reviewed articles, reports, posters), except newsletters and PowerPoint presentations |

Titles and abstracts of the retrieved records were screened by the primary reviewer based on the inclusion criteria, of which a random 10% sample were independently screened by a second reviewer (Shinkins B). Records that met the inclusion criteria at this stage, or those for which it was not possible to determine eligibility based on title and abstract, were then screened based on the full text. For those records where the full text was not available, the authors were contacted via email. The full texts of the eligible records at this stage were screened independently by the primary and second reviewers based on the inclusion criteria in Table 2-1. Inter-reviewer agreement was assessed with the Cohen's κ statistic (93). In case of any disagreements regarding a record's inclusion, a consensus-based discussion with additional researchers (Messenger MP and West RM) determined the final inclusion decision. EndNote X9 (Thompson Reuters) was used to manage the retrieved records.

### 2.2.3 Data extraction and analysis

A data extraction spreadsheet included: (i) basic descriptive information relating to the TPP (e.g. publication format, disease of interest, targeted clinical setting, funder, time-horizon); (ii) the type of information (i.e. *input sources*) used to inform each test specification listed in the TPP (e.g. expert consultation, review of the literature); (iii) the activities reported to develop the TPP (e.g. rounds of document revision, consensus meetings); and (iv) the stakeholders involved at each stage of the TPP development.

Based on the review findings, a common decision-making framework was identified across the included TPPs (*scoping*, *drafting* and *consensus-building phase* – see section 2.3.3 for details). The input sources and stakeholders invited were summarised for each decision-making phase. Where stakeholders and input sources were not explicitly reported, sources included in each TPP and descriptions of each test characteristic in the main text were reviewed.

A transparency assessment was conducted to investigate the extent to which each TPP was reproducible. Each TPP was assessed in terms of whether the adopted input sources were clearly reported, along with the activities undertaken in each decision-making phase (e.g. interviews, priority-setting exercise, Delphi-like survey), which stakeholder groups were consulted, and the funding body.

Data extraction was conducted independently by the primary reviewer and a third independent reviewer (Ayaz-Shah A). In case of disagreement, additional researchers (Shinkins B, Messenger MP and West RM) were consulted.

### 2.2.4 Test characteristic clustering

Test characteristics reported within each TPP were extracted and duplicates were removed. Test characteristics were grouped under the categories of the ACCE test evaluation framework (23). Several test characteristics, however, did not fall within any of the pre-defined ACCE categories. Three additional categories were therefore included to accommodate these additional characteristics: (i) human factors (e.g. training and education, ease of test interpretation); (ii) environmental impact; and (iii) infrastructural requirements. Table 2-2 describes each category used to cluster test characteristics across TPPs.

**Table 2-2 Overview of each test characteristic category used to cluster test characteristics within TPPs**

| Test characteristic categories | Definition |
|---|---|
| **Unmet clinical need** | Specifies the disease of interest, target population and purpose of the test (23). Within the ACCE framework, this category is labelled 'test definition'. As this category overlaps with the concept of *unmet clinical need*, it is here referred to as 'unmet clinical need' to better represent the type of information TPPs provide. |
| **Analytical performance** | Defined as the ability of a test to accurately detect and measure a given analyte (e.g. precision, trueness, analytical sensitivity and specificity, limits of detection) (27, 94). |
| **Clinical validity** | Described as "the ability of a device to yield results that are correlated with a particular clinical condition or a physiological or pathological process or state" (27). |
| **Clinical utility** | The ability of a test to affect patients' health outcomes (e.g. improvement in quality of life, longer lifespan) (30). |
| **Regulatory legitimacy** | The appropriateness of the regulatory procedures in place when introducing a new test into the market (e.g. legal requirements, product registration path) (23). |
| **Costs** | The (expected) changes in costs following the introduction of a new test into a clinical care pathway (21). |
| **Human factors** | Factors concerned with the interactions between users and devices (e.g. result document and display, ease of test interpretation, tool format and complexity) (95). |
| **Environmental impact** | Encompasses a change to the environment following an interaction with the device (96). |
| **Infrastructural requirements** | Entails "the stock of the basic facilities and equipment needed for realizing a product or providing a service" (97). |

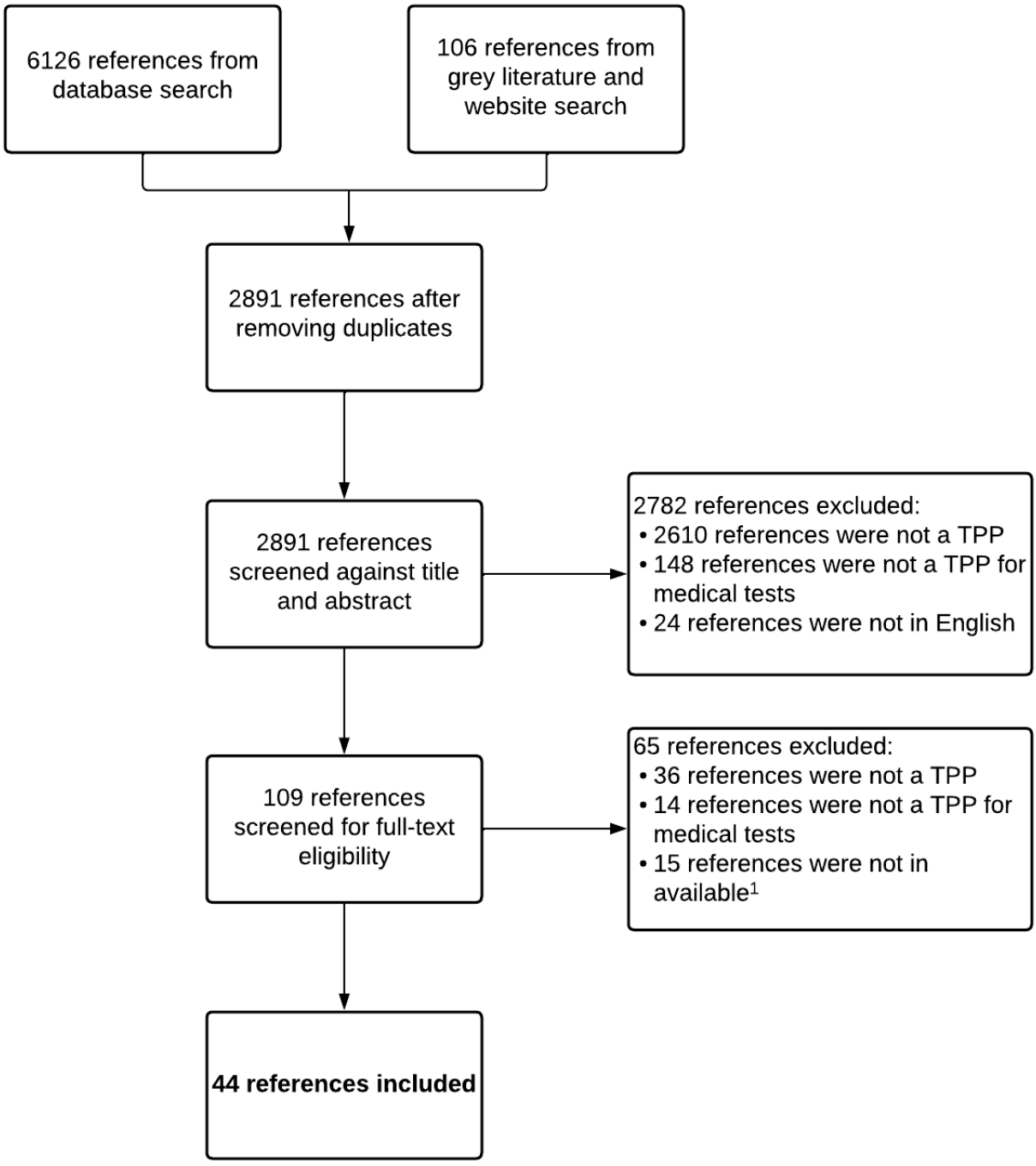
## 2.3 Results

The datasets generated and analysed during the systematic review (including the data extraction form and overview of test characteristics included in reviewed TPPs) are available in the Research Data Leeds Repository (https://doi.org/10.5518/1185).

### 2.3.1 Literature search

Full details of the literature search results are reported in Figure 2-1. From the database and online searches, 2891 records were retrieved after removing duplicates. Upon completion of the two-stage screening process, 44 TPPs were deemed eligible for inclusion (57, 60, 98-139). Twenty-eight of the 44 included records were identified from the website searches (64%), whilst 9 records were identified from the database searches (20%). Seven records were from both the database and website searches (16%). Inter-reviewer agreement between primary and secondary reviewer was high at title and abstract (κ = 96%) and full-text screening (κ = 90%). For more details, see Appendix C.

**Figure 2-1 Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) flow diagram illustrating literature search results – systematic review**




[1] For those references where full-text was not available, authors were contacted via email. For 15 references full-text was still not available.

### 2.3.2 Feature of included TPPs

The included 44 TPPs consisted of 23 reports, 16 peer-reviewed journal articles, 4 published TPP tables (a TPP without any background information or context e.g. (113)) and one conference poster. Included TPPs were dated between 2012 and 2018.

All TPPs provided guidance on developing medical tests to detect infectious diseases. Fourteen of the 44 TPPs focused on neglected tropical diseases (32%) (e.g. soil-transmitted helminths, Chagas disease, human African trypanosomiasis, schistosomiasis, trachoma, taeniasis cysticercosis), and on tests for vector-borne infections (32%) (e.g. Zika virus, Dengue fever, hepatitis C, malaria, *Escherichia coli*). Other types of infection included sexually transmitted infections (16%, n=7), respiratory infections (14%, n=6) (e.g. lower-tract respiratory infection, tuberculosis, pneumonia), Ebola virus (136), meningitis (139) and severe febrile illness (57).

Seven of the 44 TPPs were funded by Bill and Melinda Gates Foundation (16%), whereas several TPPs received funding from national agencies (99, 102, 103, 117, 120), the World Health Organisation (WHO) (57, 127, 128), or a public-private global health partnership (125). The majority of TPPs did not disclose funding sources (64%, n=28). The healthcare settings of interest were mostly low and middle-income countries.

Although the majority of TPPs outlined minimal and optimal requirements for new tests, there was a lack of agreement on the definition for ‘minimal’ and ‘optimal’ requirements. Eleven of the 44 TPPs aimed to define minimal test characteristics as the lowest acceptable specifications (25%), opposed to 6 TPPs which defined them as *must-have* characteristics (14%). One TPP defined ‘minimal’ test characteristics as those which provide a clear advantage over a comparable test (102).

Nine of the 44 TPPs reported ‘optimal’ requirements as the ideal value for test characteristics (20%), whilst 5 TPPs defined *optimal* as the values which would maximise the test’s value (11%) – without defining what ‘value’ meant in this context. One TPP reported ‘optimal’ requirements as values for a certain test characteristics that would make the test more attractive (120) – without specifying the benchmark used in this comparison. One TPP defined ‘optimal’ as the values that provide optimal ‘diagnostic effectiveness’ (102), whilst another TPP defined

them as *nice-to-have* characteristics (115), or, alternatively, “the lowest acceptable output for that characteristic for an ideal technique” (116). Two TPPs reported a single target requirement for each of the included test characteristics, rather than outlining minimal and optimal values (105, 134) .

Of the 44 TPPs identified, 7 reported the time-horizon during which the information included in the TPP was considered to be relevant for manufacturers (16%). Of those, 6 TPPs stated a time-horizon of 5 years (60, 99, 100, 104, 106, 130), whilst the remaining considered a time-horizon of 10 years (108). In three TPPs, a time-horizon was chosen to represent the timeframe within which achieving the specifications was considered feasible (60, 99, 100). In one TPP, this was based on literature reviews, surveys and consultations with industry and academic representatives (i.e. *landscape analysis* – see section 2.3.3.1) (99). In another, expected advancements in technologies and knowledge related to a certain field appeared to justify the time-horizon considered for the TPP (104).

For more details on the methodology underpinning each of the included TPPs, see Appendix C.

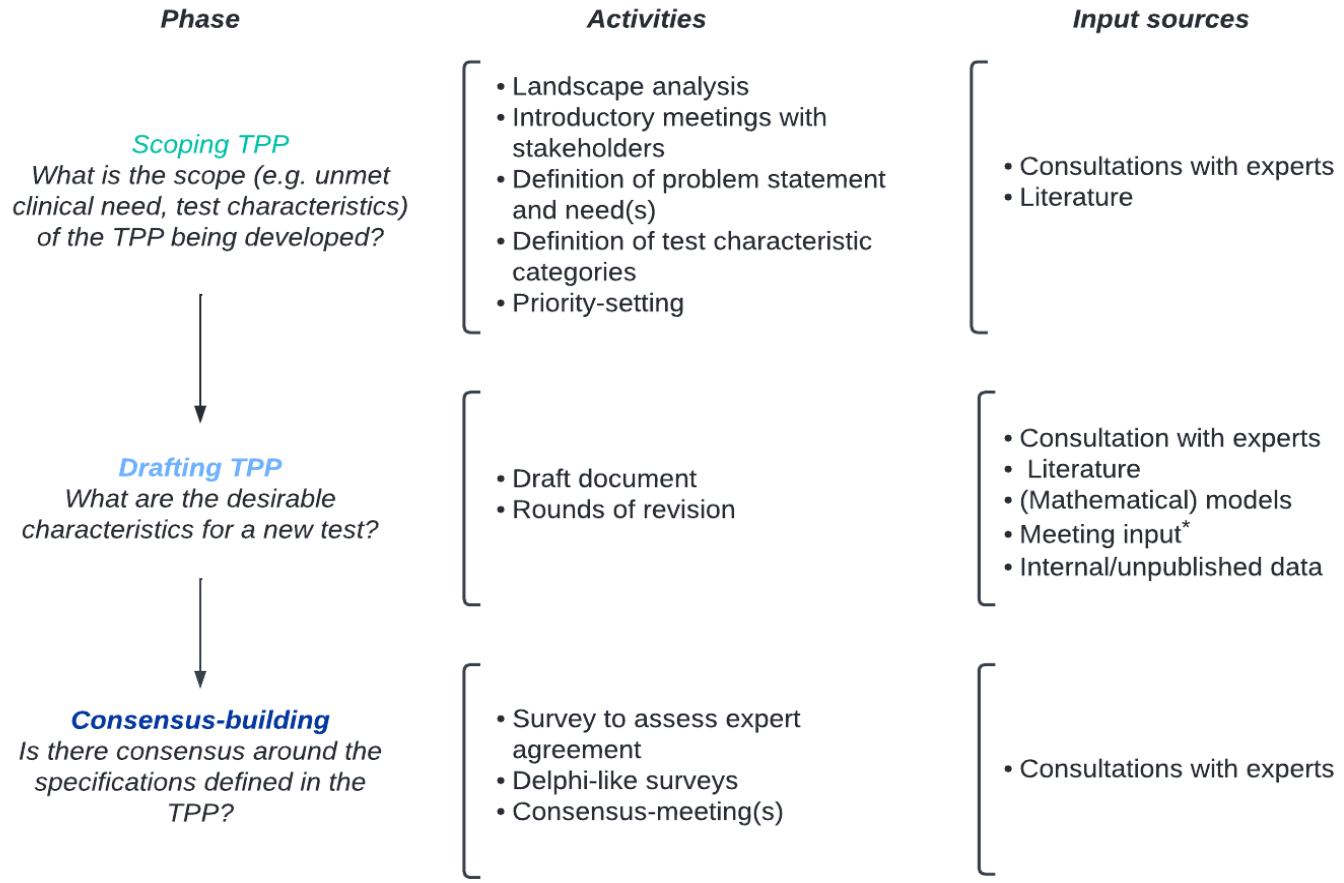
### 2.3.3 Decision-making process

A common decision-making framework, consisting of three distinct phases, was apparent across the included TPPs: scoping, drafting and consensus-building.

Figure 2-2 presents the most commonly adopted activities, and input sources used in each phase. A summary of the input sources reported to have been used during the scoping and drafting phase can be found in Table 2-3.

Figure 2-3 provides a summary of the stakeholders contributing to each phase. Several of the TPPs are excluded from Figure 2-3 as they did not report any information on input sources or stakeholder groups (112-114, 131, 133, 134, 139). The following sections describe the aim of each phase and a breakdown of the methodology (i.e. activities, input sources, and stakeholders) used within the included TPPs, where reported.

**Figure 2-2 Typical activities involved and input sources used within each decision-making phase**




* TPPs reporting 'consultations with experts' as a data source did not specify how the consultation activity took place (e.g. interviews, focus groups), whereas TPPs reporting 'meeting input' clarified that a meeting with multiple stakeholders and/or experts took place to inform TPP requirements.

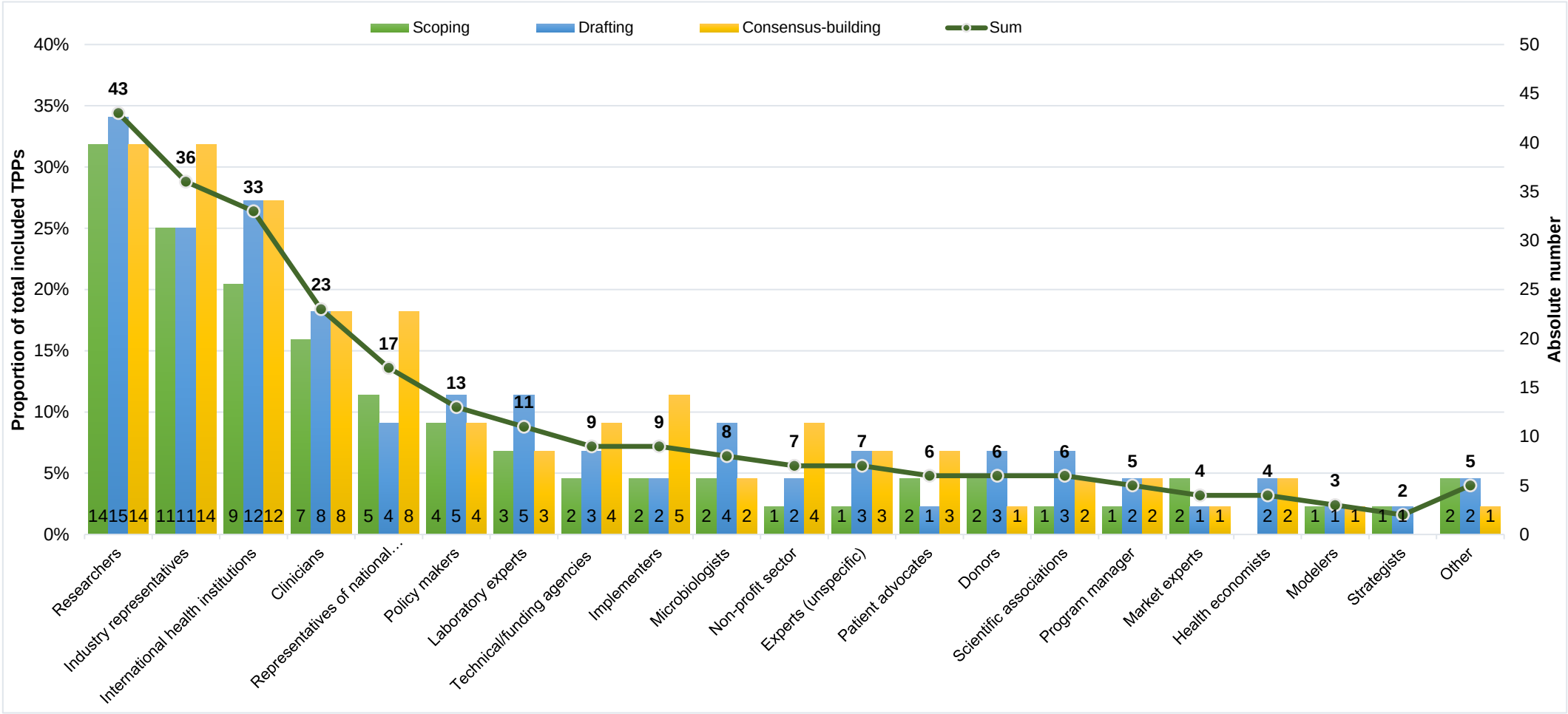



**Figure 2-3 Stakeholder groups contributing to each phase of TPP development in absolute number (n) and as a proportion of the total number of included TPPs (%)**

Note: the percentages in relation to stakeholder groups do not add up to 100% because more than one stakeholder group usually contributed to the development of a TPP

**Table 2-3 Input sources reported to be used in scoping and drafting phase**

| | Scoping | Drafting |
|---|---|---|
| | n (%)[a] | n (%)[b] |
| **TPPs reporting information on input sources considered** | 21 (48) | 33 (75) |
| **Literature** | 8 (38) | 22 (67) |
| **Expert opinion** | 15 (71) | 24 (73) |
| **Meeting input** | 4 (19) | 7 (21) |
| **Consensus-meeting input** | 0 (0) | 3 (9) |
| **Internal/unpublished data** | 0 (0) | 7 (21) |
| **Interviews/questionnaire with experts** | 4 (19) | 2 (6) |
| **Laboratory evaluations** | 0 (0) | 1 (3) |
| **Mathematical models** | 0 (0) | 8 (24) |
| **WHO clinical guidelines** | 0 (0) | 6 (18) |
| **Early usability studies** | 0 (0) | 1 (1) |
| **Industry standard** | 0 (0) | 1 (3) |
| **Field observations** | 1 (5) | 5 (15) |
| **Ethnographic interviews** | 1 (5) | 0 (0) |
| **Reports** | 2 (10) | 0 (0) |
| **WHO policies** | 1 (5) | 0 (0) |
| **Market analyses** | 1 (5) | 0 (0) |
| **Websites of developers** | 1 (5) | 0 (0) |

The percentages in relation to input sources do not add up to 100% because each TPP could list more than one source.

[a] Percentages are calculated in relation to the number of TPPs which provided information on input sources used during the scoping phase (n=21).

[b] Percentages are calculated in relation to the number of TPPs which provided information on input sources used during the drafting phase (n=33).

[c] Interviews that are conducted in a naturalistic setting to observe how study participants interact in their environment.

#### 2.3.3.1 Scoping phase methodology

Half of the TPPs described the methodology underpinning the scoping phase (n=22). The general aim of this phase is to identify the clinical problem and unmet needs that the test should address, in addition to outlining which test characteristics to include in the TPP. This phase also includes an overview of the disease area and the limitations of existing technologies.

In the scoping section, every TPP clearly defined the *purpose* of the desirable tests (i.e. screening, diagnosis, prognosis, prediction or monitoring). Most TPPs aimed to address clinical needs for diagnostic (n=39), monitoring (n=7), or screening tests (n=4)[7]. Several TPPs specified the exact technology that the new test should use (n=8) – be it either a lateral flow device (120, 122, 124), a multiplex multi-analyte cartridge assay (57), a nucleic acid amplification test (NAAT)[8] (106, 120).

Key activities reported to have been undertaken during the scoping phase included: (i) reviewing existing literature (n=6); (ii) introductory meetings with stakeholders (n=4); and (iii) *priority-setting exercises*[9] (n=5).

A minority of authors reported (n=4) that they had conducted a *landscape analysis* (22, 26, 37, 50), providing information on the disease area of interest, available diagnostic technologies, related characteristics, and limitations. Although none of the TPPs provided the definition for a 'landscape analysis', these were usually based on interviews with stakeholders and reviews of the literature. Only Toskin, I. et al. (129) reported a systematic literature review, providing the search strategy and databases searched.

Consultations with experts (n=15, 71%)[10] and literature findings (n=8, 38%)[8] were the most commonly utilised sources of information during the scoping phase (see Table 2-3 for a full breakdown). Only one type of source was considered in 15 TPPs (of which 11 was consulting experts), whilst 7 TPPs employed more than one source. Denkinger, C. et al. (99) mapped the available diagnostic tests in the

[7] Total numbers do not add up to 44 as some TPPs described required characteristics for more than one test within the same TPP document.

[8] An alternative name for NAAT is polymerase chain reaction (PCR) test.

[9] In a *priority-setting exercise*, stakeholders rank different options based on a pre-specified set of criteria (e.g. disease burden, market potential) in order to define the highest priority option.

[10] Percentages are calculated in relation to the number of TPPs which provided information on input sources used during the scoping phase (n=21).

disease area of interest and then performed a survey to gauge stakeholders' preferences. Reipold, E. et al. (127) identified the main characteristic categories (e.g. scope, performance, operational characteristics and pricing) to be included in the TPP. Five TPPs involved a priority-setting exercise during which clinical experts and stakeholders ranked each identified health need (60, 100, 106, 110, 127).

Across the TPPs included, many stakeholder groups contributed to the scoping phase, most notably researchers (n=14, 32%)[11], international public health institutions (n=12, 20%)[9], and industry representatives (n=11, 25%)[9]. Policy makers (n=4, 9%)[9], patient advocates (n=1, 5%)[9], and representatives of the non-profit sector (n=1, 2%)[9] were less frequently engaged in scoping the unmet need for a new test. For details on the stakeholders invited to the scoping phase see Figure 2-3.

#### 2.3.3.2 Drafting phase methodology

The first draft of each TPP was usually prepared by either an established working group comprising experts from different organisations (103, 108, 110, 118, 120, 129, 137) or authors of the published TPP. There were two cases where the TPP was drafted by a completely different organisation (130, 136). The TPPs were often revised several times and, in some cases, were then shortened to ensure ease of communication to different stakeholders (60, 99, 100, 106, 108).

Of the 44 included TPPs, 33 reported which input sources were considered during the drafting phase (75%) (Table 2-3). Common input sources for populating test characteristics were expert consultations (n=24, 73%) and reviews of the literature (n=22, 67%). Several also took into consideration mathematical models (n=7, 21%), WHO clinical guidelines (n=6, 18%), and 'field observations' (n=5, 15%) – although none of the identified TPPs defined what this term meant. Seven TPPs relied on unpublished or internal data to inform test performance specifications, although none of them provided any further information on the data used. Only one TPP was informed by pooled data from a systematic review (129). Industry standards, laboratory evaluations and early usability studies were the least considered types of source among the TPPs being reviewed. Twenty-six of

[11] Percentages are calculated in relation to the total of included TPPs (n=44) as per Figure 2-3.

the 44 TPPs relied on more than one type of input source at the drafting phase, whereas 7 TPPs only adopted one. Meeting inputs were the most common single source (n=3, 43%). For a full breakdown on the input sources used during the drafting phase see Table 2-3.

An investigation of which input sources informed the TPPs' analytical and diagnostic accuracy specifications was conducted, given the importance of these parameters to the performance of medical tests. For many of the TPPs, the analytical and diagnostic accuracy specifications were based on published literature (60, 99-102, 104, 106, 110, 115, 116, 118, 121-124, 127), or consultations with experts (60, 101, 103, 104, 106, 117, 127).

Seven TPPs adopted mathematic models to help define either analytical or diagnostic accuracy (60, 99, 100, 102, 106, 117, 127), including: (i) a mathematical model of malaria transmission which explored the impact of varying the diagnostic cut-off threshold in reducing malaria transmission (140); (ii) a compartment transmission model capturing changes in TB transmission and mortality (141); (iii) a model estimating the positive and negative predictive values in the context of varying disease prevalence (117); and (iv) a decision tree model measuring the effectiveness, in terms of disability-adjusted life years (DALYs), of a set of diagnostic sensitivity and specificity pairs of a hypothetical triage test for TB (87). Among the TPPs that specified quantitative minimal and desirable diagnostic accuracy values (n=36, 82%), several failed to specify which diagnostic reference standard should be used (57, 103, 109, 113, 114, 135, 137). For many of the TPPs, optimal targets for diagnostic accuracy were set to greater than 99% (100, 102, 106, 109, 110, 114, 118, 121-129, 134, 135).

Methods used to define target specifications for test prices were also investigated. The majority of TPPs included estimates of optimal and/or minimal price for new tests (n=30, 68%). Across the TPPs which reported the input sources used to inform this test characteristic (n=13), test prices were mostly informed by consultations with experts (n=5, 38%) or literature findings (n=4, 31%), followed by reports (n=2, 15%). Six TPPs used clinical studies to inform the benchmarks for test price (60, 102, 118, 121-123), including: (i) a trial-based cost-effectiveness analysis conducted in Uganda (142); (ii) an unpublished study on the end-users' willingness-to-pay for a human immunodeficiency virus (HIV) test (118); (iii) a cost-utility study conducted in Kenya (143); and (iv) an ongoing

cost and affordability study presented during the TPP consensus-meeting (60). WHO (60) also considered the findings from a published decision tree evaluating the impact of an hypothetical triage test for TB to inform the desirable test price (87).

Similar to the scoping phase, the stakeholder groups mostly engaged in drafting TPPs were researchers (n=15, 34%)[9], international public health institutions (27%, n=12) [9], industry representatives (n=11, 25%)[9] and clinicians (n=8, 18%)[9]. Implementers (n=2, 5%)[9], health economists (n=2, 5%)[9] and modelers (n=1, 2%)[9] contributed less frequently to the drafting phase (see Figure 2-3).

#### 2.3.3.3 Consensus-building phase methodology

Initial agreement with the TPP was often obtained using a survey of stakeholders (n=14). The survey either included general questions regarding stakeholders' views on the TPP (99, 102, 104, 130), or adopted a Delphi-like approach to provide an initial consensus on various aspects of the TPP (n=10). A consensus meeting with stakeholders and experts was typically held (n=11) and a revised TPP agreed upon. In some cases, an additional survey was sent to stakeholders on trade-offs between test attributes (127), or on ranking key parameters (130, 133). For two TPPs, the final TPP draft was presented to a broader stakeholder base to validate it (99, 100).

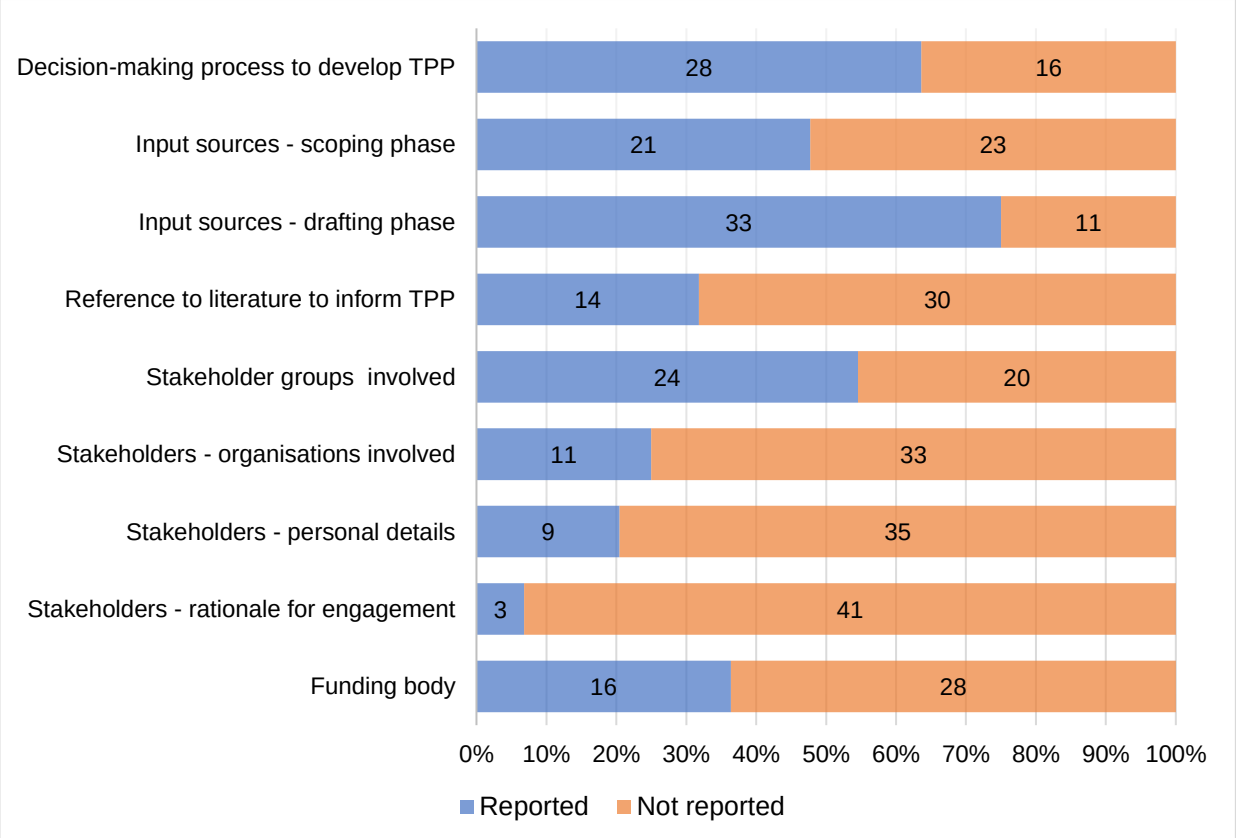
The number of participants invited to the consensus-building meetings varied (<20 participants: n=5; between 20 and 50 participants: n=7). One meeting included 100 participants (104). For several of the TPPs, the authors also took part in the consensus-meetings (60, 108, 116, 137).

Less than half of the included TPPs reported information on the activities and stakeholders invited to the consensus-building phase (n=19). A variety of stakeholder groups were invited to the consensus-building phase, ranging from researchers (n=14, 32%)[9], industry representatives (n=14, 32%)[9], international public health organisations (n=12, 27%)[9], to representatives of national disease programs (n=8, 18%) and policy-makers (n=4, 9%)[9]. The stakeholders engaged in the consensus-building phase are reported in Figure 2-3.

### 2.3.4 Methodological transparency

The transparency of the TPPs in terms of reporting their methods was also assessed (see Figure 2-4). Over a quarter of the included TPPs did not state the decision-making process behind TPP development (n=16, 36%). Further to this, many failed to report which sources were considered to inform the scoping phase (n=23, 52%) and to populate the TPP during the drafting phase (n=11, 25%). Just under half did not disclose which stakeholders were involved in the development of the TPP (n=20, 45%). Specifically, the name of the organisations stakeholders were affiliated with was only reported in 11 TPPs (25%), whilst 9 TPPs mentioned personal details of each stakeholder (20%). Most TPPs failed to explain why certain stakeholders were involved in the development of the TPP (n=41, 93%) – except three studies which stated that stakeholders were engaged in virtue of their expertise and experience in the field of interest (103, 116, 137). Sources of funding were often not reported (n=28, 64%). In a minority of TPPs the methodology was clearly reported (60, 103, 106, 116). For further details on the transparency assessment of each of the included TPPs see Appendix C.

**Figure 2-4 Transparency in methods reporting across the included TPPs (n=44)**



### 2.3.5 Test characteristics of included in TPPs

After removing duplicates, 140 different test characteristics were reported across the included TPPs. Features which did not represent test characteristics have been excluded from this analysis, such as factors relating specifically to the disease in question, rather than the test. Figure 2-5 shows the test characteristics most frequently reported (a full list is available in Appendix D). A full list of test characteristics is available in Appendix D, and the 'Test Characteristic Clustering' Excel spreadsheet available at the Research Data Leeds Repository (https://doi.org/10.5518/1185).

**Figure 2-5 Test characteristics frequently reported in all TPPs (n=44) sorted by categories**

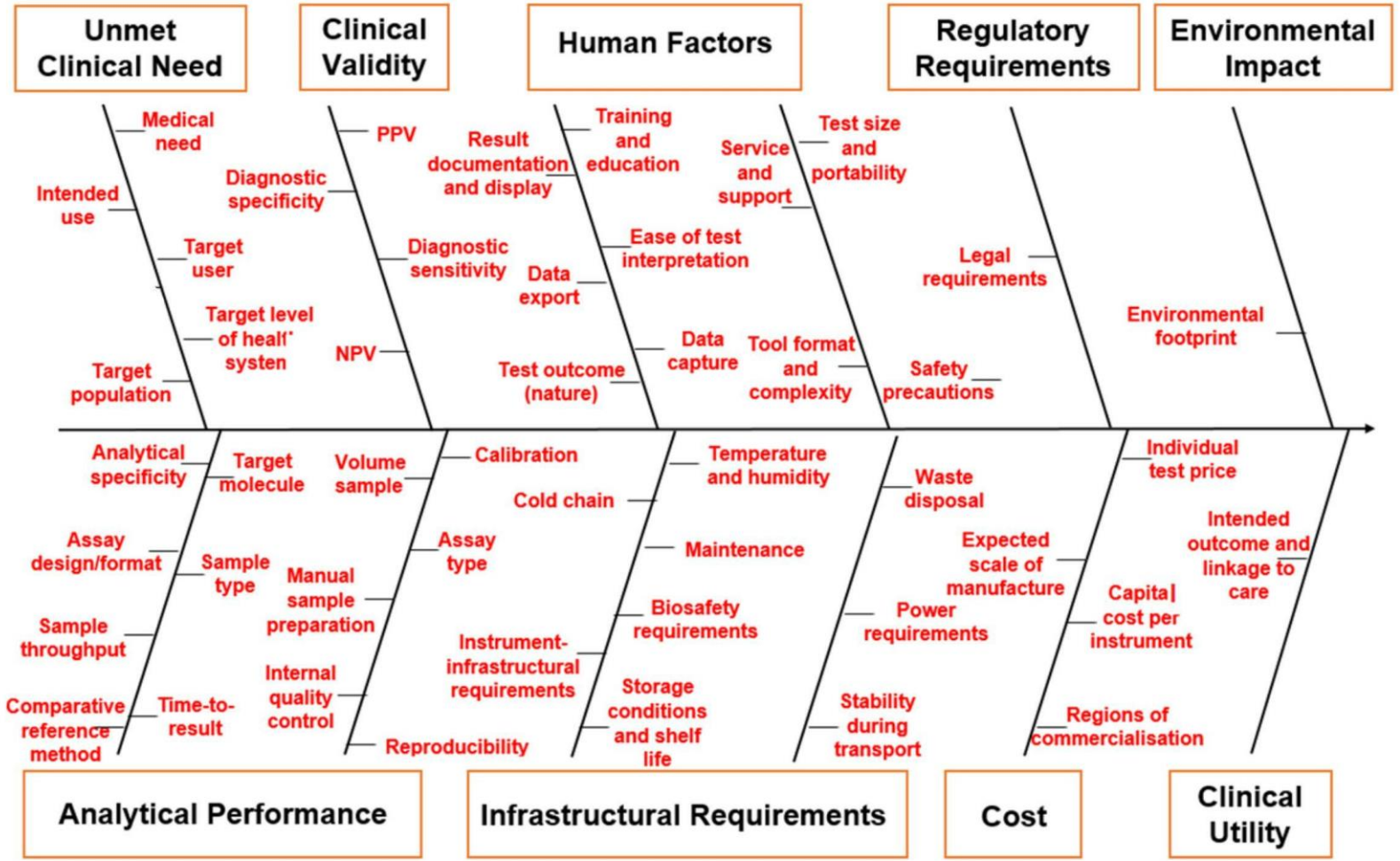
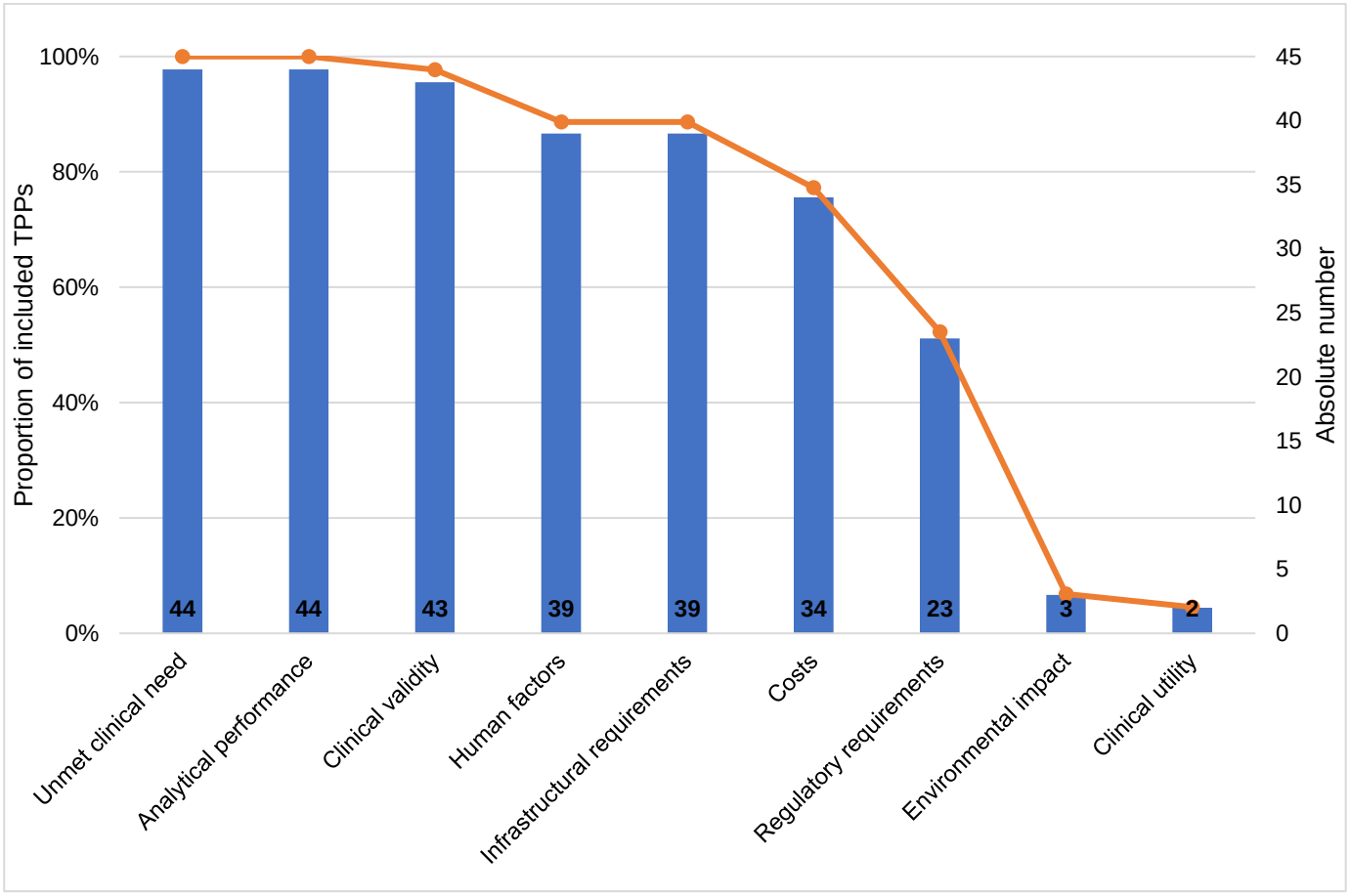


Figure 2-6 depicts which characteristic categories were reported in the included TPPs. Details on unmet clinical need, analytical performance and clinical validity appeared to be consistently reported, whereas regulatory impact, environmental footprint and clinical utility were less frequently considered.

**Figure 2-6 Test characteristic categories included in TPPs, in absolute number (n) and as a proportion (%)**



## 2.4 Discussion

### 2.4.1 Review findings

The findings of this systematic review describe the methods currently used to develop TPPs for medical tests. Forty-four TPPs dated between 2012 and 2018 were identified, all of which focused on infectious diseases.

A shared framework for the development of TPPs was identified, comprising of three core phases: scoping, drafting and consensus building (as described in Figure 2-2). The scoping phase aims to identify the focus of the TPP, both in terms of the unmet clinical need in question and which test characteristics should be included in the TPP. Consultations with experts and information from the literature were typically considered during the scoping phase. The drafting phase entails developing and revising the TPP, and was mainly based on consultations with experts, literature findings, meetings and available data (e.g. internal and unpublished data). The final phase, consensus-building, aims to gather stakeholder's views on the TPP through surveys and consensus-meetings. Across all phases of TPP development, the most frequently involved stakeholder groups were academic researchers, industry representatives, international public organizations, clinicians and members of disease-specific programs.

TPPs for new tests to date provide manufacturers with necessary specifications mostly related to test analytical performance (e.g. volume sample required) and clinical validity (e.g. predictive values, diagnostic accuracy), followed by infrastructural requirements (e.g. storage conditions), human factors (e.g. tool format and complexity) and costs. The environmental impact, and clinical utility and cost-effectiveness of tests were rarely considered. Extensive variability in the terminology for test characteristics was also found, with several alternative terms used to refer to common test performance specifications. This, in turn, might hinder the interpretability and applicability of TPPs for test developers and researchers in this area, especially if the use of an alternative terminology is not justified.

There was a general lack of transparency and consistency in reporting the methods underlying TPPs. This makes it difficult both to appraise the validity of the specifications reported in the TPPs, and to ascertain whether the specifications are expected to be generalisable to other settings. The lack of clear

reporting in this regard further makes it challenging for the specifications within TPPs to be updated in the future.

Nevertheless, it is unclear if manufacturers develop tests strictly in line with TPPs, or whether there are any factors which make this infeasible or challenging. A recent survey indicated that test developers found TPPs for POCTs for sexually transmitted diseases generally useful, but they suggested highlighting which test performance specifications are perceived as most important within TPPs, as it is often considered extremely challenging to achieve every presented desirable target (144).

This review shows that publicly available TPPs for tests to date have primarily been developed for global health applications, with the main funding organisations being WHO, UNICEF and Bill and Melinda Gates Foundation. The sole focus on infectious diseases may be explained by the remit of these global organisations. For example, WHO included HIV, neglected tropical diseases, TB and malaria as priority diseases (145). Further to this, WHO established the 'R&D Blueprint', which aims to promote R&D activities (tests, vaccines, medicines) during epidemics and pandemics (146). After having identified the pathogen to target first, TPPs are usually commissioned to guide the development process of new healthcare products which will address the high-priority pathogen (146). Similarly, the Bill and Melinda Gates Foundation supports the development of healthcare interventions to reduce the burden of the leading causes of mortality and morbidity in low- and middle-income countries, with a particular interest on malaria, TB and neglected tropical diseases (147).

The predominance of global health organisations – such as the WHO – in the development of (publicly available) TPPs might be explained by the following reasons:

- ***driver of innovation*** **in countries most in need** – one of the strategic goals of the WHO is to scale and sustain innovation in countries and areas most in need (148). Their role is therefore pivotal in signposting areas of greatest unmet clinical need in order to stimulate R&D process of new health technologies in countries where there is less financial incentive for innovation. A potential application of this could be found in 'R&D Blueprint' initiative discussed above.

- **role of *standard setter*** – one of the key roles of the WHO is to improve global health via the definition of common norms and standards with which health products and interventions should comply (149, 150). Global health organisations are intergovernmental organisations which are usually considered above individual nation-states' and private enterprises' interests. They might therefore be regarded as an impartial authority in charge of setting priorities in relation to health conditions and defining performance benchmarks for new health products.
- **access to extensive resources** – developing a well-executed TPP is a multi-disciplinary exercise which involves several consultations with key health leaders and stakeholders. Global health organisations might have greater advocacy and communication resources, as well as an increased capacity to contact and retain long-term collaborations with key opinion leaders and stakeholders across different settings – especially in the context of low- and middle- income countries (149).

### 2.4.2 Identified limitations of current TPP methodology

In reviewing current methodology for developing TPPs for medical tests, key areas for improvement were identified. These concerned: (i) a reliance on subjective data to inform test requirements; (ii) a lack of explicit consideration of clinical utility when defining test characteristics; and (iii) a focus on price rather than cost-effectiveness when considering costs. The following sections describe each of these limitations in turn and the implications on the development process of TPPs for medical tests.

#### Subjectivity of Input Sources

Expert judgment and selected evidence identified in published literature were the main sources of information used for defining desirable and acceptable test characteristics. Systematic reviews of the literature, where extensive and reproducible database searches are carried out and where the quality of relevant studies is appraised, were not conducted to identify relevant evidence at the scoping and drafting phase in any of the identified TPPs – except one (129). A possible reason for not conducting systematic reviews of the literature might lie in an expected paucity of evidence in the literature at the onset of TPP development, due to the early nature of these documents – although it is difficult to verify this claim based on the reported methods. Nevertheless, it is unclear

how the TPP developers have identified the relevant literature to inform test specifications as search strategies were seldom reported – regardless of the type of literature review. This is likely to introduce bias and subjectivity in terms of the evidence used to underpin test characteristic recommendations.

Although expert judgement is undoubtedly useful, relying solely upon this information source has limitations, particularly when the aim is to derive quantitative estimates. How humans make probability judgements is strongly affected by many heuristics and systematic biases (e.g. anchoring, availability, overconfidence and hindsight bias[12]) (151). Specifically, previous literature has found a poor understanding of test accuracy among healthcare professionals (153) as diagnostic sensitivity and specificity are often misinterpreted and mistaken for predictive values (153). It is of interest that, among the included TPPs, only Denkinger, C. et al. (100) relied on the findings from a decision-analytic model, which was used to back-calculate the minimal specifications for diagnostic accuracy for a new triage TB tests based on DALYs (87). As discussed in Chapter 1, decision-analytic modelling offers a framework of analysis within which a full range of relevant evidence (e.g. literature findings, datasets, expert opinion), and uncertainty around that evidence, can be synthesised in order to provide a simplified specification of the decision problem (61).

Additionally, the quality of expert elicitation heavily relies on the process of expert selection, as it is important to choose experts with good subject knowledge. Only 4 TPPs described how this selection process took place, and therefore how representative the expert judgments were might be questioned. Furthermore, many TPPs reported literature as a source for informing TPPs, although fewer than half of the TPPs cited the references considered. This lack of transparency hinders the quality and credibility of sources on which TPPs are based. If methods

[12] *Anchoring* is a cognitive bias that leads individuals to rely exclusively on the first piece of information acquired in relation to a certain topic, or pre-existing information, when making a decision (151).

*Availability bias* reflects the tendency to rely on an immediate piece of information, rather than seeking for the most appropriate source of information, when evaluating a decision (151).

*Overconfidence* is the tendency of a person to overestimate and inflate their own skills and knowledge about a certain topic when making a decision – thereby the person would not seek additional evidence (152).

*Hindsight bias* reflects the tendency to exaggerate the predictability of a certain event only once that event has occurred (151).

and sources are not clearly reported, this will make it challenging for the TPP to be updated in light of additional evidence.

**Oversight of Clinical Utility**

Only two of the TPPs reported desirable characteristics relating to the clinical utility of the test. This is unsurprising given that the majority of research efforts in the field of test evaluation focus on generating evidence on analytical and clinical performance (e.g. diagnostic accuracy) (27). A highly analytically- or clinically-accurate test, however, does not necessarily mean that the test will improve patient health, as factors relating to clinical decision-making and effectiveness of patient management and treatment strategies must also be considered to make that assessment (154).

In the case of a new test, estimating the clinical utility is particularly complicated given the uncertainty around the mechanisms by which the test will impact on patient outcomes (29), and common difficulties in obtaining long-term observational or RCT data (30, 155). Nevertheless, decision-analytic modelling may be of use in this context as a means of estimating the clinical utility of a test. With decision-analytic modelling, it possible to link the accuracy of a test with evidence of treatment effectiveness to capture the downstream consequences of testing on patient health (30, 155). This helps to estimate the efficacy of a test-treatment strategy in terms of intermediate and final outcomes that are relevant for decision-makers (29). Although less than a quarter of the TPPs used mathematical models to inform test specifications, none of them were used to derive performance benchmarks for new tests in this context.

**Focus on price rather than cost-effectiveness**

Although the minimum and optimal price of the tests featured in many of the TPPs, none of these were driven by the trade-off between the overall cost implications of implementing the test (e.g. cost of clinical management, treatment and long-term clinical events) and the associated patient benefits. The exception was one TPP that used the findings from a model-based cost-effectiveness analysis to estimate the minimum acceptable test cost (60). A small subset of the included TPPs suggested conducting further cost-benefit analyses and ‘health economics studies’ to inform the required price for new tests (110, 127). It is important to consider the cost of the new test in the context of the complete

benefits and harms that the test may provide, as well as the broader costs of patient management and downstream clinical events. For example, a new test may be relatively expensive but may also improve patient health to an extent that justifies this additional cost. Conversely, a new test may be relatively cheap but offer no improvements in patient health and therefore even the marginal increase in cost is not justified.

Failure to consider cost-effectiveness implications when setting the price of a (existing or new) diagnostic test might hinder the reimbursement and adoption process of the test into clinical practice. Decision-makers increasingly demand evidence that a new test improves patient health and is cost-effective rather than solely evidence of its analytical and clinical validity (21). Specifically, many HTA bodies in Europe, Australia and North America consider clinical utility, cost and cost-effectiveness in relation to the target population, in addition to analytical performance and clinical validity when assessing new molecular diagnostic tests (156). In addition, the new European IVDR is set to demand evidence on the clinical utility of tests, in order for test producers to obtain a CE mark (27). As such, the inclusion of clinical utility and cost-effectiveness considerations in TPPs would bring these documents more in line with the wider regulatory and reimbursement landscape, and thus increase the utility of TPPs in the R&D pipeline.

Conducting cost-effectiveness analysis at early R&D stages of new tests can help manufacturers to avoid significant investments in tests that do not have the potential to be cost-effective (29). Whilst economic evaluation is usually conducted at the later stages of the evaluation pathway, EEE and cost-effectiveness analysis are increasingly performed at early development stages as a means of directing *stop/go* decisions and informing the optimal trajectory of future research (71, 72).

### 2.4.3 Study limitations

Since this study is a systematic review of publicly available literature, a key limitation is that the search strategy will have inevitably missed confidential or unpublished TPPs developed in-house by test manufacturers. Although the results of the online searches did not identify any companies stating that they had developed TPPs for medical tests, it is unlikely that such information would be publicly available on company websites due to commercial interests. On the

contrary, it is more likely for global health organisations to publish their TPPs as a means of steering technological advancements in areas of greatest clinical needs.

Despite TPPs for any medical test being searched, all of the reviewed TPPs focused on tests for infectious diseases. It is unclear therefore how applicable the findings from this systematic review will be to TPPs conducted for other conditions (e.g. non-communicable diseases), settings (e.g. high-income countries) or for industry-led TPPs. Nevertheless, it may be argued that TPPs conducted by large established bodies, such as WHO and the Bill & Melinda Gates Foundation, would be expected to represent current best practice methods in this area. Further research, attempting to gather data on the methodology adopted by manufacturers developing in-house confidential TPPs, could be conducted to test this hypothesis.

### 2.4.4 Recent applications of TPPs

In response to the COVID-19 pandemic, the use of TPPs as a mechanism to steer technology R&D development for new diagnostics has begun to be more broadly recognised in international and national organisations. This section describes the most notable TPPs published since the completion of this review.

In 2020, the WHO issued several TPPs for healthcare interventions to tackle the COVID-19 pandemic (157), including vaccines (158) and diagnostics (159); while the Foundation for Innovative New Diagnostics (FIND) has been gathering information on the commercially available or in-development diagnostic tests for COVID-19. Given the high volume of diagnostic tests for COVID-19 currently marketed (over a thousand, as of October 2020) (160), having pre-specified performance benchmarks outlined in TPPs is expected to support the development and assessment process of new diagnostics for COVID-19 (161).

In addition, TPPs are increasingly of interest to national policy makers seeking to accelerate and de-risk the R&D process of innovative diagnostics. In the UK, the Medicines & Healthcare products Regulatory Agency (MHRA) developed a series of TPPs for COVID-19 tests, aimed at assisting manufacturers to design and deliver tests to support the UK's testing strategy and inform test validation, procurement and regulatory decision-making (162). The performance benchmarks and technical specifications listed in the MHRA TPPs lie at the core

of the UK National technical validation process for manufacturers of COVID-19 tests (163). Characteristics of new tests under development are initially reviewed and compared against the technical specifications included within the MHRA TPPs. This, in turn, informs a *go/no go* decision as only tests meeting the criteria listed in the TPPs start the formal technical validation and in-service evaluation. This is the first example of TPPs being used by national policy makers to inform and de-risk the validation and procurement process of new tests.

In addition to the MHRA, NICE in the UK have also recently shown interest in the use of TPPs to guide technology development and evaluation. In 2021, an early economic model of COVID-19 POCTs was published as part of a NICE-commissioned project (164), which aimed to assess the costs and benefits of having a hypothetical POCT for diagnosing COVID-19 which met the desirable performance specifications as presented in the MHRA TPPs (162). As the MHRA TPP had already been developed, this model-based early economic model did not aim to inform the performance specifications set out in the TPP, but rather to identify the biggest drivers of cost-effectiveness changes (165). Despite the novelty of this modelling approach, the results of this analysis were evaluated under the traditional late-stage economic evaluation perspective, rather than as exploratory results aiming to support the development process of new tests.

Although the COVID-19 pandemic has put a spotlight on the value of TPPs to provide clinically meaningful and evidence-based performance specifications for new diagnostics, their application has still almost exclusively been limited to infectious diseases. There have been recent calls to extend the application of TPPs to other disease areas – in particular, in their recent report on early cancer detection and diagnosis (ED&D), Cancer Research UK (CRUK) has called for the creation of a national body to "map, define and conduct an evidence assessment of the ED&D product pipeline in the UK and globally, and to set Target Product Profiles (TPPs) (e.g. desired performance characteristics for new tests to work towards), with a UK government and NHS commitment to rapidly adopt tests which meet these TPPs" (45). Only time will tell if this initiative is undertaken, but clearly this is a positive step towards a wider understanding and utilisation of TPPs – especially in the context of main funding bodies in high-income countries.

### 2.4.5 Future research

Although there is evidence of a common underlying development framework for TPPs, this review highlights that there is considerable variability in TPP practices. In particular: (i) there is a lack of harmony in terminology used to refer to key test performance characteristics; (ii) the methods employed to draft TPPs are inconsistent and difficult to assess due to lack in transparency in methodology reporting; and (iii) there is no clear justification as to why certain test characteristics are included. Guidance on best practice methods for developing TPPs for medical tests would be highly beneficial, and timely given the recent wider interest in adopting TPPs (see section 2.4.4). Akin to the USA FDA guidance on TPPs for drugs (outlined in Chapter 1), a guidance document could be developed for TPPs for medical tests summarising the purpose, attributes of TPPs and which test characteristics should be included.

Future research should also focus on how to identify unmet clinical needs underpinning a certain disease area. Monaghan, P. et al. (18) developed a valuable checklist for identifying biomarkers based on literature findings and consultations with experts. This checklist could be pertinent for the scoping phase underlying TPP development, although this would need further validation in this specific context. Based on this checklist, an online survey to identify unmet clinical needs is developed and applied to the thesis case study – see Chapter 4 for details on the survey development and results.

Most importantly, research is required to understand how to better incorporate the assessment of desirable clinical utility and cost-effectiveness of innovative tests into TPPs. In particular, investigation to explore if and how care pathway analysis and early economic modelling could be integrated into the development of TPPs is required. It is expected that care pathway analysis would provide clarity on the mechanisms by which a test could impact on downstream patient outcomes; whilst early economic modelling could be used to define desirable values for certain test characteristics (e.g. test price, diagnostic accuracy) based on cost-effectiveness considerations (67). The potential benefits early economic modelling and care pathway analysis could bring to TPP methodology development, as well as potential implications, are further explored in section 2.5.

## 2.5 Integrating early economic evaluation methodology within TPP development

In this section, the stated hypothesis that EEE could strengthen the methodological rigour of the typical TPP development (presented in section 2.3.3) is discussed. Based on the methodological limitations underpinning the typical TPP development process (discussed in section 2.4.2), key possible advantages that EEE methodology could afford across the four core activities of TPP development – *scoping* (see section 2.5.1), *drafting* and *consensus-building* (section 2.5.2), and *updating* (section 2.5.3) – are described. The key arguments raised in this section are summarised in Figure 2-7.

**Figure 2-7 Overview of expected benefits that EEE can bring to each TPP development activity, alongside relevant methods**

| *TPP development activity* | *EEE expected advantages* | *EEE methods* |
|---|---|---|
| ***Scoping the unmet clinical need for new tests***<br>*What is the current clinical pathway and how does a new test fit into it?* | Pathway mapping to understand standard care, and to define new test role, purpose and potential impact | • Care pathway analysis<br>• Scenario analysis (structural uncertainty) |
| ***Drafting and agreeing test specifications***<br>*What are the desirable characteristics for a new test and what factors have most impact on them?* | Exploring dependencies between test specifications and features of the clinical context | • Deterministic sensitivity analysis<br>• Scenario analysis (parameter uncertainty) |
| | Estimating ranges for quantitative test characteristics | • Headroom analysis<br>• Threshold analysis<br>• Sensitivity analysis |
| | Objective derivation of test specifications | Transparent reporting of model analysis and results |
| ***Updating TPP specifications***<br>*How do test specifications change if new evidence occurs?* | Incorporating new evidence | Transparent reporting of model structure, parameters, analysis, results |
| | Highlighting key areas for future research | • Value of information analysis<br>• Sensitivity analysis |

### 2.5.1 Scoping: pathway mapping to define test role, purpose and potential Impact

The scoping phase defines the focus of the TPP in terms of the unmet clinical need being addressed, and which test characteristics should be included (1). Based on current practice, it is unclear if any formal pathway analysis and/or assessment of the test's placement and impact is typically undertaken at this stage (1); this risks overlooking key mechanisms of impact on patient outcomes.

The first step in economic modelling is to map the care pathway in which the new technology will sit. This typically involves reviewing clinical guidelines, and consulting clinicians, patients and carers on their individual experiences of the care pathway. The result is a clear schematic of the existing clinical pathway, detailing the main activities and events which may occur for the patient population under consideration. This ensures that the processes involved in standard care are clearly understood and also forces test developers to specify the test role (i.e. whether the test is a replacement, triage or add-on) and purpose (e.g. screening, diagnosis, prognosis, prediction or monitoring) within the specified pathway. This is crucial in the context of TPP development, since selecting a different test role and/or purpose can result in significantly different test requirements (77). For example, in the management of patients with COVID-19, molecular tests can be used for screening, triage, diagnosis, monitoring and prognosis of patients across a wide variety of clinical settings (e.g. home, doctors surgery, emergency department) and applied to populations with differing disease severity and prevalence of infection. Where the optimal placement of a new test is unclear, additional comparators can be added to the EEE to explore alternative options (77).

Pathway mapping further helps to identify the possible downstream consequences (harms and benefits) of a new test, including expected impacts on: decision-making (e.g. change in treatment decisions), patient health outcomes, clinical workflow (e.g. time-to-treatment) and economic outcomes (166). This is vital for TPPs, since the expected harms and benefits associated with a test should directly influence which test specifications are included within the TPP. For EEEs, this step further helps to identify which test properties should be captured and varied within the economic model (e.g. diagnostic accuracy, test price, turnaround-time). Although pathway mapping is a core process of

economic modelling, it should be noted that this activity can be conducted independently of any formal economic evaluation.

### 2.5.2 Drafting and agreeing test specifications

The TPP drafting phase (see section 2.3.3.2) estimates the specifications for new tests, whilst consensus-building aims to attain consensus on those requirements through stakeholder surveys and consensus-meetings (section 2.3.3.3). Typically, test specifications are presented at two levels (*desirable* and *acceptable*) without explicitly considering the dependencies between different test specifications (e.g. the dependence between diagnostic sensitivity and specificity), or other factors related to the clinical context (e.g. the required sensitivity and specificity will be dependent on disease prevalence) that could jointly impact the overall utility of a given testing strategy (1). This apparent failure to consider parameter dependencies risks over- or under-estimating test requirements.

The following subsections present how EEE can: (i) allow exploration of dependencies between test specifications and features of the clinical context; (ii) facilitate optimisation of quantitative test specifications based on a WTP threshold; and (iii) increase the objectivity of data informing TPP requirements.

*Exploring dependencies between test specifications and features of the clinical context*

Desirable ranges for test specifications may vary when considering different factors underlying the care pathway or other test properties. It might therefore be helpful to distinguish upfront between:

- **properties inherent to the test** - test properties (e.g. diagnostic accuracy, test turnaround time) that could have a direct/indirect effect on the utility of the test; and
- **factors relating to the clinical context** - aspects exogenous to the test that could influence how the test impacts upon relevant outcomes (e.g. prevalence of disease, natural disease progression, efficacy and cost of treatment).

With EEE, by synthesising multiple test attributes and factors relating to the clinical context within a single modelling framework, the dependencies between different parameters can be effectively captured and explored. Univariate and

bivariate sensitivity analyses can be used to explore the impact of one or more parameter(s) on the clinical utility or cost-effectiveness (75, 82-86). Scenario analyses can also be used to explore the impact of a group of parameter changes that represent a specific clinical scenario (75).

*Estimating ranges for quantitative test characteristics*

Outside the context of TPPs, EEE has previously been used to identify acceptable ranges for test characteristics or components – see Chapter 1. In the absence of evidence for a new test, hypothetical values for test properties can be assigned and, for each specification or combinations of specifications, the downstream costs and benefits can be computed. Based on a specified WTP threshold, it is possible to then derive the maximum costs and minimum specifications for a new test to be cost-effective (72, 84).

Based on the published examples of early economic models for diagnostics outlined in Chapter 1, key relevant methods which may be useful in this phase include: headroom analysis, threshold analysis, and sensitivity analysis (Chapter 1).

*Objective derivation of test specifications*

The early stages of technology development are characterised by high uncertainty, making it challenging for experts to form accurate judgments around the desirable specifications a new test should possess.

EEE methodology facilitates the generation of more objective data, which can be combined with subjective evidence (i.e. clinical and stakeholder judgment) to inform TPP specifications. EEE results can also support the elicitation of expert opinion, with clinical experts being asked to review and discuss the modelling findings. Presenting the results of EEE modelling within consensus-meetings could further help to narrow down the ranges for performance benchmarks a new test is required to meet in order to be clinically and cost-effective.

Were EEEs to be integrated into TPP development, the methods, model code and analysis underpinning the TPP should be clearly reported and made freely available, in line with international modelling guidelines (167). This would provide a more transparent and objective approach to setting performance benchmarks while, in turn, allowing others to inspect what data have informed TPP development.

### 2.5.3 Updating TPP specifications

Although TPPs are dynamic documents which should be updated as new evidence is found (49), no formal or standardised approach for updating TPP specifications is currently available (1). This activity is vital for TPPs as a means of ensuring that test development is driven by accurate, up-to-date information. Expected advantages EEE can bring to this core activity are presented below.

*Incorporating new evidence*

EEEs can be designed to be sufficiently flexible to allow for the model to be iteratively updated as evidence is accrued (29). This means that different TPPs could be developed based on the level of evidence available: early TPPs would focus on defining the unmet clinical need and key requirements for a new test to address that need; whilst later iterations could identify more precise test specifications.

Abel, L. et al. (75), for example, published their model code to allow others to examine and explore additional questions. For TPPs, this approach would help to ensure that performance specifications could be validated and updated with each iteration of the TPP, thus further increasing the methodological rigor underpinning TPPs.

*Highlighting areas for future research*

Findings from sensitivity analyses could be used to communicate areas of greatest uncertainty requiring further research. Alongside this, when more evidence and information about the test is available (i.e. when parameter distributions reflect uncertainties rather than *unknowns*), VOI analysis may be useful to assess the value of gathering more information on model parameters to reduce current decision uncertainty (77, 85). To calculate the VOI, it is necessary to express the uncertainty surrounding test characteristics (e.g. price, diagnostic accuracy) numerically. At early TPP development stages, these test features are unknown at the outset and they are derived using the WTP threshold. Any assumption made around these key parameters would therefore drive the results of the VOI analysis leading to potentially misleading conclusions.

### 2.5.4 Further considerations

Section 2.5 discussed how key methodological limitations with current TPPs for medical tests could be addressed by integrating EEE into the TPP development process. Clearly the EEE methods discussed could be independently applied to the TPP development based on the level of evidence available and project aim, be it care pathway analysis while scoping the unmet clinical need, or decision-analytic modelling and threshold analysis were TPP developers in the process of deriving performance specifications, or sensitivity analysis while updating TPP specifications.

Nevertheless, there are potential limitations and complexities with integrating EEE into TPP development that should be considered. For example, conducting a wide range of sensitivity analyses with EEE risks overwhelming TPP developers with information, making the estimation of clearly defined test performance requirements challenging. Such analyses should always be kept within the confines of clinical plausibility, and clinical experts should be consulted to support the identification of clinically relevant scenarios. Given that uncertainty in the technology development process is unavoidable, particularly in the early stages of R&D, it is suggested that the value in such analyses lies in explicitly communicating those key areas of uncertainty, identifying trade-offs between test characteristics, and highlighting priorities for future research.

Economic evaluation also requires the adoption of a particular jurisdiction (e.g. UK NHS), which dictates what country-specific clinical pathways and costs are included in the model, and what pre-specified decision criterion is used to inform the analysis (e.g. the NICE WTP threshold per QALY). In case a particular decision criterion is not based on cost-effectiveness, the EEE methods discussed in section 2.3.3.2 may still be useful to explore scenarios around clinical utility or cost alone. The applicability of modelling findings to other jurisdictions is therefore often limited, and multiple model versions may be required to derive test specifications across different jurisdictions. This is true of any formal evaluation of downstream clinical or cost outcomes however (where country-specific clinical pathways and costs will come in to play), and is not specific to EEE.

Whilst EEE is expected to strengthen the methodological rigour of TPP development, it is recommended to regard this approach as an adjunct rather than a replacement for existing TPP development methodology. Expert input and consensus-building discussions with relevant stakeholders will be key to ensuring that EEEs capture the nuances of the clinical context, and that the results are clinically meaningful. Ultimately this will ensure the efficient development of tests which provide greater utility for both patients and healthcare services.

## 2.6 Chapter summary

- Based on 44 TPPs for medical tests, this review identified a common TPP development framework comprising of three decision-making phases: (i) scoping; (ii) drafting and (iii) consensus-building (summarised in Figure 2-2).
- Within this framework, key limitations were identified: (i) a heavy reliance on expert opinion to inform test specifications; (ii) poor transparency in methodological reporting; (iii) a lack of explicit consideration of clinical utility; and (iv) an oversight of cost-effectiveness considerations.
- Based on these limitations, the potential benefits of integrating EEE methodology into the development process for TPPs were discussed (summarised in Figure 2-7). Using EEE methods to inform TPPs is expected to provide a more objective, evidence-based and transparent approach to defining test specifications, while also capturing the key aspects underpinning the care pathway within which a new test will sit.

The remainder of this thesis is focused in applying the EEE methodology to a case study of a rapid diagnostic tests for CDI (Chapter 5 and Chapter 6). The clinical context of the case study is first provided in Chapter 3. Chapter 4 then presents the findings of an online survey of UK healthcare professionals which aimed to scope unmet clinical needs for new CDI diagnostics.

# Chapter 3
# Case study clinical background

## 3.1 Chapter outline

In Chapter 2, a systematic review of TPPs for medical tests was conducted and several methodological limitations were identified. As discussed, some of these limitations could be addressed by integrating EEE into the TPP development process. The suggested EEE methods discussed in Chapter 2 are applied to a case study on new rapid diagnostic tests for CDI in Chapter 5 and Chapter 6.

CDI was selected as a case study for this thesis for the following reasons: (i) the development of a TPP for new diagnostic tests for CDI is part of a MRC-funded programme grant (as stated in section 1.6.1); (ii) there is currently a paucity of evidence on the perceived unmet clinical needs hindering the diagnosis of CDI, thereby there was potential utility in eliciting views from relevant stakeholders via an online survey; and (iii) selection of this case study offered a unique opportunity to explore and measure the impact of a new rapid and more accurate diagnostic test for CDI on key resources used in the clinical care pathway (e.g. single rooms where to isolate patients confirmed with CDI in order to minimise the infection spread) via the development of a resource-constrained DES model (further described in Chapter 5).

This chapter provides the clinical background to the thesis case study. The clinical context is first presented (section 3.2), followed by an outline of the laboratory diagnosis of CDI (section 3.3). An overview of diagnostic strategies for CDI is then presented (section 3.4), focusing on diagnostic tests detecting a single target (i.e. *standalone tests*) (section 3.4.1), and combinations of diagnostic tests (i.e. *testing algorithms*) (section 3.4.2). The clinical pathway for those suspected and confirmed with CDI based on the UK clinical guidelines is then described (section 3.5).

A glossary of key terms used in this chapter is provided in Appendix E.

## 3.2 Clinical context

Clostridioides difficile (*C. difficile*) – formerly called *Clostridium* – is an anaerobic, gram-positive, spore-forming, toxin-producing pathogen which causes CDI. The following sections describe the biology (see section 3.2.1), typical symptoms (section 3.2.2), disease spread and infection-control measures (section 3.2.3) and healthcare costs associated with CDI (section 3.2.4).

### 3.2.1 Biology

CDI is one of the leading causes of healthcare-associated infections, resulting in approximately 120,000 cases and 3,700 deaths across Europe every year (168). In England, 13,286 hospital-associated cases (24 cases per 100,000 individuals) were reported in the 2017/2018 financial year (169).

Prior antibiotic treatment is one of the main risk factors for CDI, followed by use of gastric acid suppressants, older age (>65 years) and prolonged hospitalisation (170, 171). Exposure to antibiotics increases the risk of CDI by disrupting the normal bowel flora and allowing for the opportunistic proliferation of the pathogen (172). Upon administration of antibiotic treatment, some bacteria – including *C. difficile* – have increasingly become resistant to certain antibiotics, causing suboptimal clinical outcomes for patients and treatment failures (173).

### 3.2.2 Symptoms

Typical symptoms of CDI can range from mild diarrhoea to colitis, and toxic megacolon, which may result in death (174-176). Most strains of *C. difficile* produce two toxins, toxin A (TcdA) and toxin B (TcdB), which are considered the leading cause of CDI symptoms (177). Table 3-1 outlines the common symptoms of CDI, sorted by disease severity.

**Table 3-1 CDI symptoms sorted by disease severity based on UK clinical guidelines for CDI (178)**

| Severity of disease | Symptoms |
|---|---|
| **Mild** | – Normal white blood cell count (WCC)<br>– Stool frequency < 3 per day; and<br>– Stool consistency equal to type 5-7 on Bristol Stool Scale[a] |
| **Moderate** | – WCC higher than normal values but below 15 x $10^9$ per litre; and<br>– Stool frequency equal to 3 to 5 per day |
| **Severe** | – WCC higher than 15 x $10^9$ per litre; or<br>– An acute rising serum creatinine (i.e.>50% increase above baseline); or<br>– Fever (i.e. >38.5°C); or<br>– Evidence of severe colitis |
| **Life-threatening** | – Hypotension; or<br>– Partial or complete ileus; or<br>– Toxic megacolon; or<br>– Radiological evidence of severe disease |

[a] Bristol Stool Chart is a medical tool aid used to visually classify faeces into seven groups depending on stool consistency, ranging from constipation (type 1) to diarrhoea (type 7) (179).

After receiving antibiotic treatment for CDI – be it either metronidazole, vancomycin or fidaxomicin depending on disease severity and local cost-effectiveness considerations (180) – symptoms of diarrhoea typically resolve within one or two weeks (181). A proportion of individuals confirmed with CDI (20% to 30%), however, might experience disease recurrence after receiving initial antibiotic treatment for CDI (182-184).

### 3.2.3 Infection spread and infection control measures

The pathogenesis of CDI starts with the ingestion of *C. difficile* vegetative organism or spores (185). Patients infected with CDI shed spores in stools which, in turn, leads to the contamination of their skin, clothing, bed linen and other common environmental surfaces (e.g. toilet, basin, tables) (186). This increases the risk of in-hospital transmission of the *C. difficile* pathogen between patients, healthcare workers and visitors.

Many infection-control measures are therefore in place to curb the spread of CDI within hospitals – including antimicrobial stewardship, single room isolation of symptomatic patients, hand-washing policies and environmental disinfection (187, 188). Diagnostic tests for CDI are also part of multifaceted strategies aiming

to control the spread within hospitals (187). In the UK, inpatients suspected with CDI are typically isolated in single rooms while awaiting test results to prevent in-hospital transmission, until confirmation has been received that they have non-infectious diarrhoea (181).

### 3.2.4 Healthcare costs

In addition to a prolonged hospital length of stay (LOS) (189-191), high morbidity and mortality (192), CDI is associated with increased healthcare costs and resource use (193, 194). In the UK for example, two observational studies estimated an average healthcare cost of £6,294 to £12,710 over a LOS of 15 to 17 days (194, 195). After hospital bed and drug costs, total laboratory costs represent a key cost driver, ranging between £296 (95% CI: £223-£369) (194), and £304 (IQR £142-£404) (195) per diagnosis for a cohort of 45 and 64 patients, respectively.

## 3.3 Laboratory diagnosis of CDI

The initial diagnosis of CDI is challenging due to: (1) uncertainty regarding who to test for CDI; (2) the number of different targets used for detection of CDI; and (3) the lack of agreement regarding the diagnostic reference test. The following sections describe each of these issues separately.

### 3.3.1 Uncertainties regarding who to suspect of CDI

Only individuals presenting with loose (diarrheal) stools should be tested for the presence of CDI (196). The presence of diarrhoea, however, is common among hospitalised patients (12-32%) (197), which can be due to: (i) other health conditions – both infectious or non-infectious causes (198); or (ii) medical treatments. This means that having diarrhoea is not sufficient to suspect with confidence that a patient is infected with CDI.

Distinguishing between antibiotic-related diarrhoea (not infectious) and *C. difficile*-mediated diarrhoea is challenging as there may be no difference in clinical symptoms between these two conditions (199). Among the infectious causes of hospital-associated diarrhoea, *C. difficile* is the most common organism, followed by other pathogens such as *Salmonella*, *Campylobacter*, *Shigella* and *Escherichia coli* (198). There are also many different non-infectious causes of diarrhoea, including certain treatments (e.g. laxatives, oncologic

therapy, radiotherapy, antibiotics), or underlying conditions (e.g. inflammatory bowel disease, irritable bowel syndrome, gastrointestinal [GI] neoplasia, celiac disease, anxiety) (198).

### 3.3.2 Multiple targets to detect

Another source of complexity underlying the diagnosis of CDI lies in the extensive range of targets for CDI (200) (e.g. *C. difficil*e toxins in stools, *C. difficile* organism, and toxigenic strains of *C. difficile*), or a combination of these targets. How each target correlates with the clinical diseases of CDI differs; and detecting one analyte is not necessarily sufficient to confirm the diagnosis of CDI in a patient with diarrhoea.

For example, detecting the *C. difficile* organism in stools does not confirm CDI, since the *C. difficile* strains could be either: (i) *non-toxigenic* which are not regarded as pathogenic; or (i) *toxigenic* (i.e. having the ability to produce toxins due to the presence of a toxin-encoding gene or active toxins) (201). This distinction is particularly relevant given that high levels of asymptomatic colonisation have been reported among healthy individuals (3%-7%) (202, 203), patients upon hospital admission (4%-15%) (202, 204) and elderly patients in care homes (50%) (205, 206). Further testing is therefore required to assess whether the *C. difficile* strains carry the toxin genes and/or are producing toxins.

Detecting solely the presence of toxigenic strains in stools without checking for active toxin production is also not sufficient to confirm the presence of CDI (206, 207). The expression of *C. difficile* toxins is considered the main driver of symptoms related to CDI, especially toxin B (208) – thereby leading to the conclusion that CDI is a toxin-mediated disease and that detection of toxins in stools is a pre-requisite for diagnosing CDI (209-211). It is therefore crucial to test for active toxin production in the context of a toxigenic strain to ultimately confirm the diagnosis of CDI. However, there is an ongoing debate surrounding the clinical significance of detecting a toxigenic strain in stool without any free detectable toxins (198).

How the various CDI tests available may be used to inform the diagnosis of CDI is discussed further in section 3.4.

### 3.3.3 Lack of agreement on the reference testing standard

Two testing options are considered diagnostic reference standards for the diagnosis of CDI, namely the cell cytotoxicity neutralisation assay (CCNA) and toxigenic culture (TC) (200). There is an ongoing debate, however, around which reference standard to use when evaluating new and existing testing options for CDI (212, 213). The choice of the reference test to use is dependent on the target of interest (213, 214). Where the aim is to detect free toxins in the sample, CCNA appears to be the appropriate standard method whereas, if the aim is to detect a toxigenic strain, TC seems to be a more suitable reference method (213).

Some studies have reported a higher diagnostic sensitivity of TC compared to CCNA (208, 215), concluding that TC detects a third more cases compared to CCNA (216, 217). However, positive results with CCNA have been reported to correlate better with clinical outcome and mortality than TC (218), meaning it may be a better diagnostic tool for CDI than TC. As these tests detect different targets and there are no standardised procedures for running these tests, it is difficult to compare these diagnostic reference standards head-to-head, especially across different laboratories (213).

The lack of agreement on the optimal diagnostic reference standard hinders the development and assessment of new diagnostic tests (219), since evaluating tests detecting different targets will not facilitate the estimation of the true accuracy of a test. This issue is also likely to impact upon the current understanding of the epidemiology of CDI (206). Differences in incidence and prevalence rates across countries might, in part, reflect differences in diagnostic practices rather than the true frequency of *C. difficile* infected patients (206).

## 3.4 Diagnostic strategies for CDI

There are two main groups of diagnostic strategies for CDI, including: (1) *standalone tests* – diagnostic tests detecting a single target (see section 3.4.1); and (2) *testing algorithms* – a combination of two or more standalone tests (section 3.4.2).

### 3.4.1 Standalone tests for CDI

Different testing methods, detecting different targets, are available for diagnosing CDI (see Figure 3-1 for more details). These tests can be grouped depending on the target being detected, including:

- **toxins in stools** – CCNA and enzyme immunoassays (EIA) testing kits for toxin A and B detect free toxins in stools which are produced by the organism but not bound to it. While CCNA has high diagnostic specificity (200, 220), the long turnaround time (24 to 72 hours), and high labour intensity of this test hinder its widespread adoption as a routine testing option. Although EIA testing kits for toxin A/B is a popular testing option for CDI (221), European and UK clinical guidelines recommend against their use as standalone tests as their diagnostic accuracy is considered suboptimal (222, 223).
- ***C. difficile*** **organism** – EIA testing kits for glutamate dehydrogenase (GDH)[13] and culture confirm the presence of the pathogen in stools. While EIAs detecting GDH are recommended as initial screening tests due to their high negative predictive value (200, 224-227), these tests suffer low diagnostic specificity for CDI (where active production of toxin is required) (227-229), as well as low positive predictive value (222, 230, 231). In the context of a positive result for GDH, it is therefore necessary to assess if there is also active toxin production in stools, whereas a negative GDH result can ruled out the diagnosis of CDI with confidence.
- **toxigenic *C. difficile* strain** – TC and NAAT detect the presence of a toxigenic strain of *C. difficile*. TC assesses the ability of *C. difficile* strains to be toxigenic *in vitro*, which may not have actually occurred *in vivo*. TC therefore does not differentiate between asymptomatic colonisation and active infection (208, 226, 232). While NAATs are commonly adopted as a testing option for CDI, they detect the presence of toxin-encoding genes which may not automatically translate into an actual production of toxins (i.e. low diagnostic specificity).

Standalone tests for detecting CDI present different advantages and limitations (see Figure 3-1 for further detail), with none of the current standalone tests being able to accurately and/or quickly identify patients with CDI (222). To overcome the limitations associated with the available standalone tests for CDI, different

[13] GDH is a metabolic enzyme through which the *C. difficile* pathogen manages oxidative stress arising from the immune response (216, 224).

testing algorithms have been designed to provide a more rapid and accurate diagnosis (222, 223) (described in section 3.4.2).

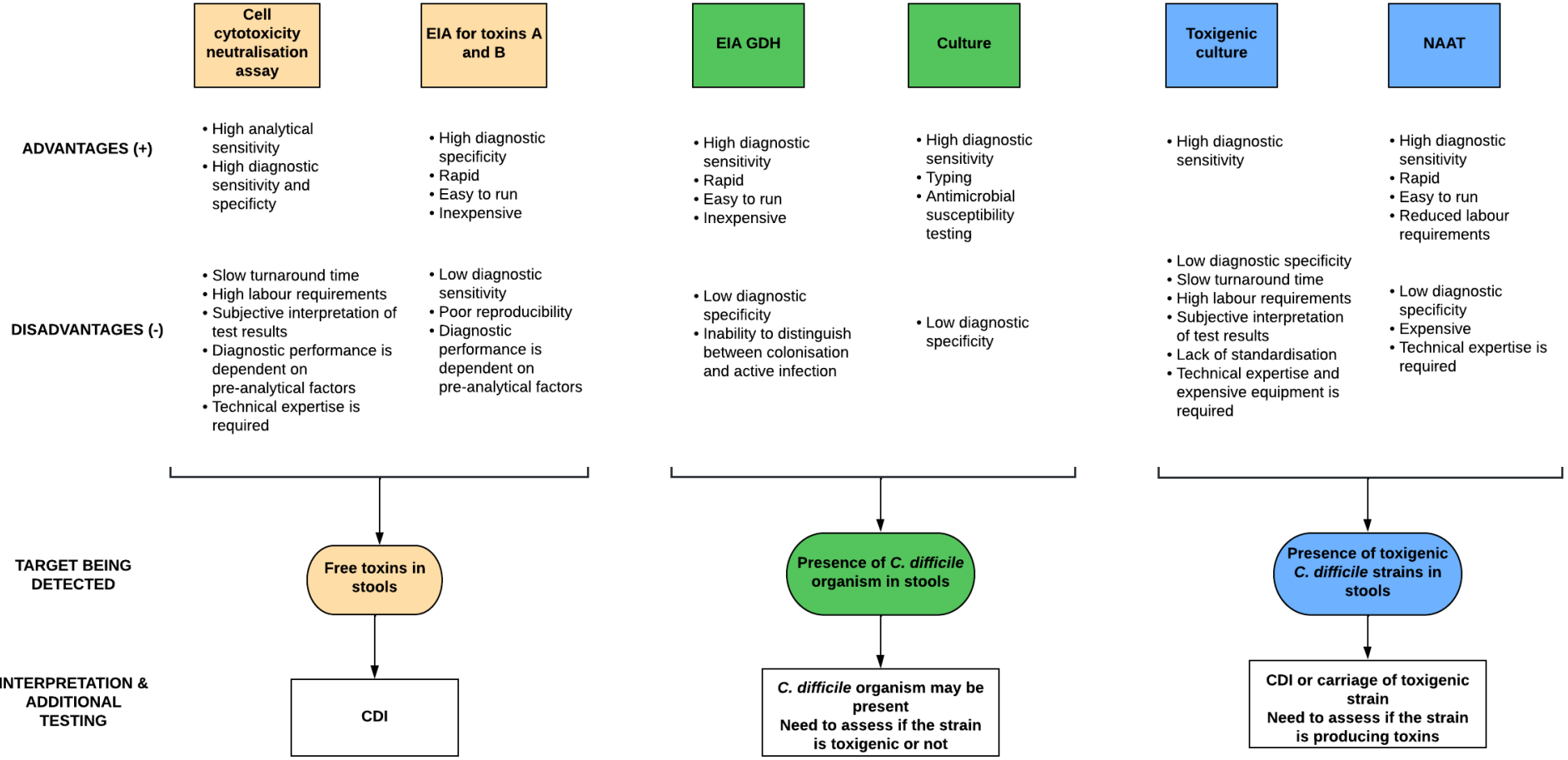



**Figure 3-1 Advantages and disadvantages of the standalone diagnostic tests for CDI, and underlying interpretation of test results sorted by target being detected. Standalone tests detecting toxins and *C. difficile* organism in stools are highlighted in orange (left two columns) and green (middle two columns), respectively. Standalone tests detecting toxigenic strains are highlighted in blue (right two columns). Figure adapted from Gateau, C. et al. (200).**

### 3.4.2 Testing algorithms for CDI

Following the limitations associated with each standalone test for CDI, combining two or more standalone tests within a testing algorithm is currently regarded as the best strategy for diagnosing CDI (222, 223). The overarching goal of testing algorithms is to achieve a higher sensitivity compared to EIAs for toxins A/B, whilst getting a quicker diagnosis with a cheaper test than NAAT and reducing TC workload (233).

Several factors inform the choice of testing algorithm – such as available testing equipment, possibility for batch testing, laboratory's expertise and financial constraints (213). The aim of the first test run as part of an algorithm is to rule out uninfected patients who can be managed as CDI negative cases (206). Testing options with high diagnostic sensitivity are therefore preferred as an initial screening test (213, 222, 223) – be it either EIA for GDH or NAAT (200, 206, 220, 228, 234). The choice of which screening option to use depends on the financial constraints of a laboratory, and the volume of stool samples run per day (200, 206, 220, 228, 234). The aim of the second step of a testing algorithm is to rule out the presence of toxins in stools, and thereby a second test with high specificity should be adopted – such as EIA for toxins A/B or CCNA (222).

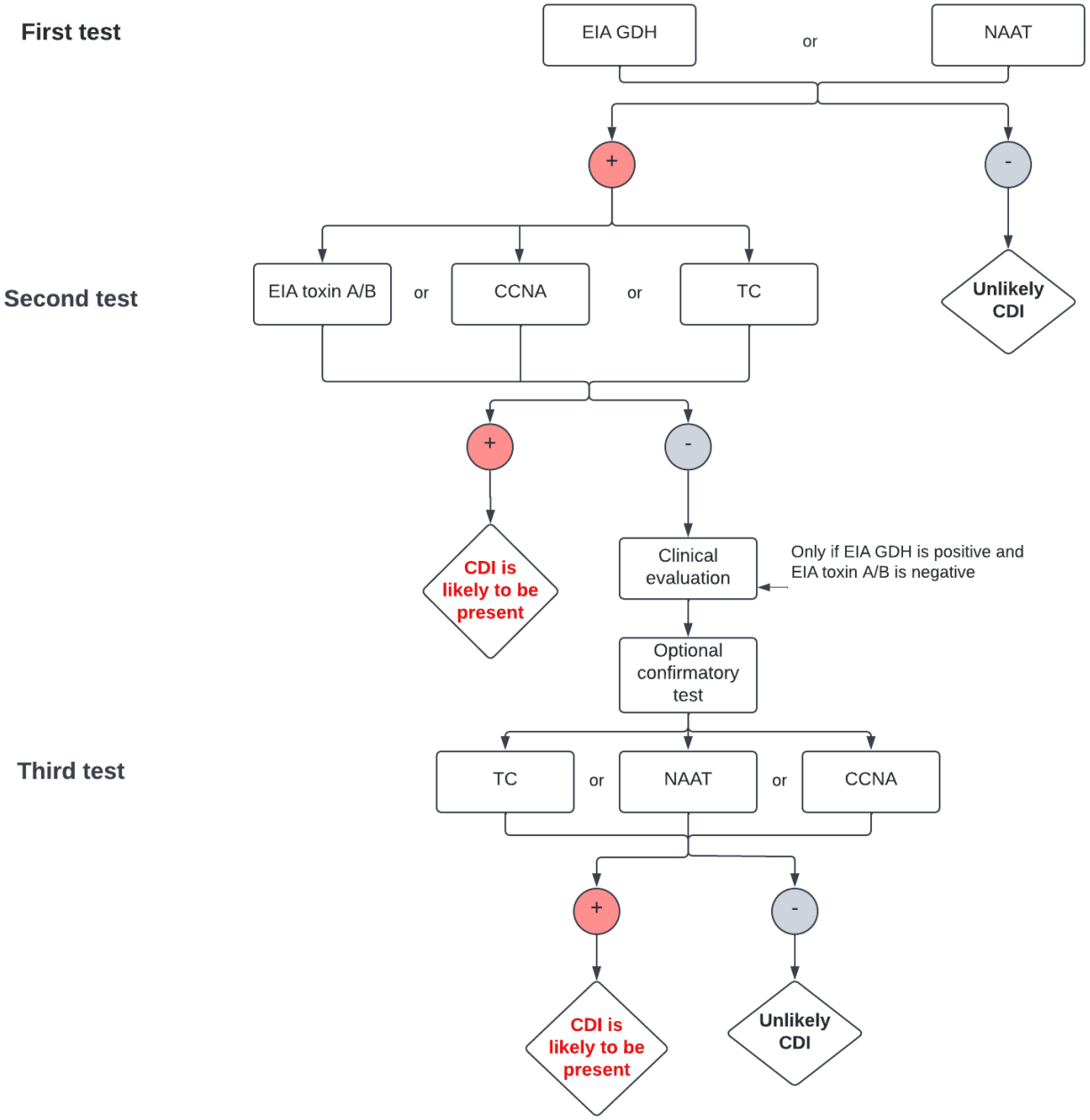
There are some variations in how testing algorithms are run (see Figure 3-2), including:

- **Order of testing options** – two or more testing options could be run either: (i) one after the other (i.e. sequential testing algorithms); or (ii) at the same time (i.e. simultaneous testing algorithms).
- **Need for confirmatory testing** – if the first and second test yield discordant results, a third confirmatory test may be run to rule in or rule out the diagnosis of CDI with confidence – such as NAAT or CCNA (200).

A key advantage of testing algorithms lies in a reduced percentage of false-positive (FP) cases as a result of improved diagnostic accuracy (213). Testing algorithms also appear to be the most cost-effective diagnostic strategy for CDI for the following reasons (216):

- **improved test-treatment pathway** – they help to reduce inappropriate antibiotic treatment for CDI, as well as the related downstream consequences on patients' health (235);
- **optimisation of the testing workflow** – depending on the testing algorithm, running a highly sensitive and cheap GDH screening test rules out negative patients with results being quickly communicated to the clinical team (236), while only positive samples are then tested with a second test;
- **discourage repeat testing for CDI** – the lack of confidence in test results may lead clinicians to repeat testing (222). Testing algorithms thus help to minimise the re-submission of multiple stool samples thereby leading to lower testing costs (206).

**Figure 3-2 Schematic of the available testing algorithms for diagnosing CDI**



## 3.5 Clinical pathway for patients suspected and confirmed with CDI

This section outlines the typical clinical pathway for patients suspected and confirmed with CDI based on the UK Department of Health and Social Care (DH) and Public Health England (PHE) clinical guidelines (see Figure 3-4).

A core recommendation in the clinical management of patients suspected with infectious diarrhoea is for clinicians (e.g. consultants, nurses) to follow the mnemonic protocol *Suspect Isolate Gloves Hand washing Test* (SIGHT) (178) – as presented in Figure 3-3.

The following subsections describe the main clinical recommendations in regards to: (i) suspicion of CDI (see section 3.5.1); (ii) testing protocol for CDI (section 3.5.2); (iii) confirmation of CDI diagnosis (section 3.5.3); (iv) antibiotic treatment options for CDI (section 3.5.4); and (v) infection-control measures (section 3.5.5).

| | |
|---|---|
| **S** | Suspect that a case may be infective where there is no clear alternative cause for diarrhoea |
| **I** | Isolate the patient and consult with the infection control team (ICT) while determining the cause of the diarrhoea |
| **G** | Gloves and aprons must be used for all contacts with the patient and their environment |
| **H** | Hand washing with soap and water should be carried out before and after each contact with the patient and the patient's environment |
| **T** | Test the stool for toxin, by sending a specimen immediately |

**Figure 3-3 SIGHT protocol according to UK DH clinical guidelines (178)**

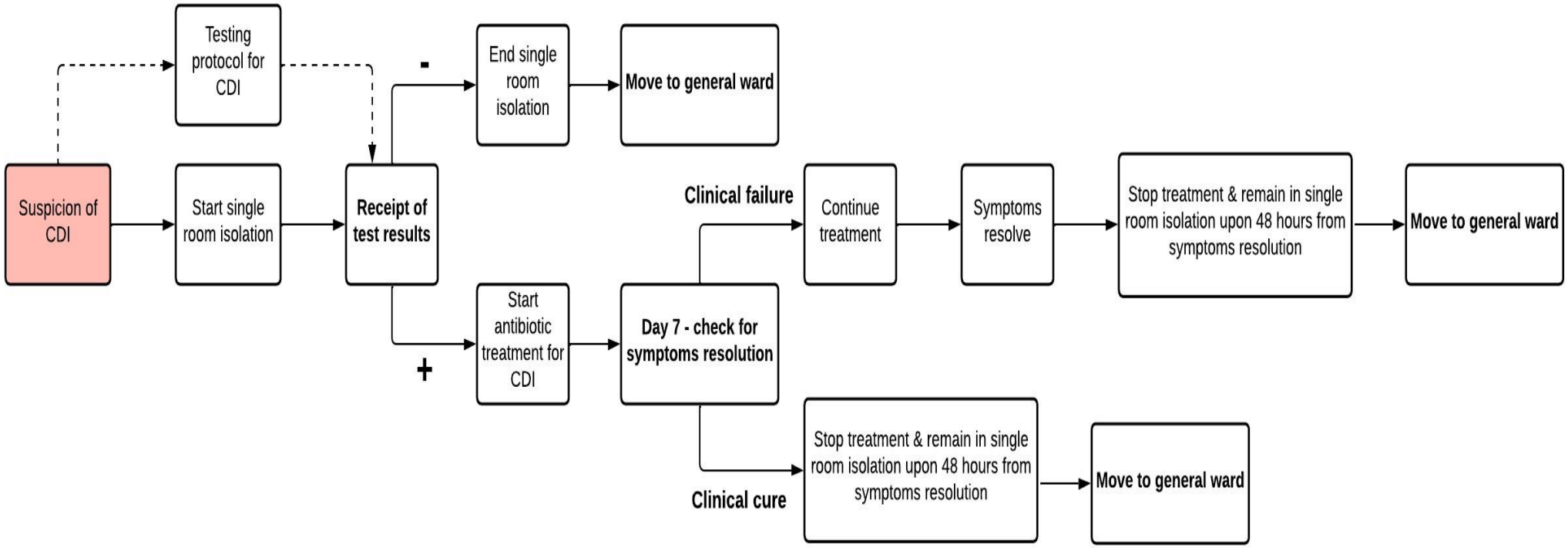



**Figure 3-4 Simplified schematic of the clinical care pathway for patients suspected and confirmed with CDI based on the UK DH and PHE clinical guidelines**

### 3.5.1 Suspicion of CDI

The UK DH adopts the Bristol Stool Chart to define a patient having diarrhoea (Bristol Stool Chart types 5-7) (223).

DH guidelines recommend that all diarrhoeal samples that are not attributable to other conditions (e.g. inflammatory colitis, overflow) or therapies (e.g. laxatives, enteral feeding) should be collected and tested for CDI as quickly as possible (223). Specifically, diarrhoeal samples from the following patient populations should be tested for CDI (223), including:

- hospital patients aged ≥ 2 years;
- all community patients aged ≥ 65 years; and
- community patients aged < 65 years whenever it is considered clinically relevant.

If a patient suspected with CDI presents severe symptoms (see Table 3-1), clinicians should pre-emptively start antibiotic treatment before test results for CDI are back (henceforth referred as *empirical treatment*) (223). Where infectious diarrhoea is suspected, patients should be isolated, ideally in a single room with a self-contained toilet and a basin, to minimise the risk of in-ward transmission of *C. difficile* or other infectious pathogens (223).

### 3.5.2 Testing protocol for CDI

DH guidelines explicitly discourage the use of EIA testing kits for toxins A/B as standalone tests for the diagnosis of CDI. The use of a two-step testing algorithm for diagnosing CDI is recommended instead (222), screening first with GDH EIA (or NAAT), followed by a sensitive EIA testing kit for toxin A/B (or, alternatively, CCNA), in the context of a positive first test result. If a sample tests negative for GDH or toxigenic strains, there is no need for further testing for toxins production in stools. If test results are discordant results (e.g. GDH positive and toxin negative), it is acceptable to test samples with an additional confirmatory test (e.g. NAAT).

In addition, it is not recommended to conduct repeat testing after a positive sample for the following 28 days due to continued shedding of the organism even after primary antibiotic regimen – except where symptoms disappear then reappear, and clinicians need to confirm the presence of recurrent CDI (178). Re-submission of samples which have already tested negative is also discouraged

as these samples can be accurately classified as non-CDI – except in the context of a strong ongoing clinical suspicion for CDI (178).

### 3.5.3 Receipt of test result

Since laboratory testing alone cannot accurately differentiate between asymptomatic colonisation from active infection (237), DH guidelines state that diagnosis of CDI is ultimately a clinical decision which should be based on a combined assessment of clinical symptoms and laboratory test results (223). Given that no standalone test or testing algorithm is regarded as perfectly accurate, antibiotic treatment decisions are ultimately clinical decisions which should not be guided solely by laboratory results. For example, if test results are negative, it might still be appropriate to administer treatment for CDI in cases of high clinical suspicion.

Table 3-2 gives an overview of the possible diagnoses and interpretation of test results based on both EIA GDH (or NAAT) and EIA toxins A/B results, according to the DH guidelines (223). Upon confirmation of non-infectious diarrhoea, patients should be released from single room isolation and clinicians should consider other causes of infectious diarrhoea. If a patient's test results confirm CDI, all non-*C. difficile* specific antibiotic treatments should be stopped to allow the normal intestinal flora to recover. In addition, antibiotic treatment for CDI should start upon confirmation of CDI (178) (see section 3.5.4 for more details) – unless clinicians had already provided *empirical treatment* in the context of severe symptoms of CDI.

**Table 3-2 Interpretation and action taken based on test results, according to DH guidelines (223)**

| | Combined result | Interpretation | Action taken |
|---|---|---|---|
| **GDH (or NAAT) -ve/ toxin A/B -ve** | • No *C. difficile* strain detected (GDH-ve); or no toxigenic strain detected (NAAT-ve)<br>• No free toxins detected | CDI is unlikely to be present | • Consider other causes of infectious diarrhoea<br>• End single room isolation |
| **GDH (or NAAT) +ve/toxin A/B +ve** | • *C. difficile* strain detected (GDH+ve); or toxigenic strain detected (NAAT+ve)<br>• Free toxins detected (toxin A and B +ve) | CDI is likely to be present | • Continue (or start) single room isolation<br>• Continue (or start) antibiotic treatment regimen for CDI |
| **GDH (or NAAT) -ve/ toxin A/ B +ve** | • No *C. difficile* strain detected (GDH -ve); or no toxigenic strain detected (NAAT-ve)<br>• Free toxins detected (toxin A/B +ve) | Rare, patients need to be re-tested | • Retest for CDI |
| **GDH (or NAAT) +ve/ toxin A/ B –ve (at least one)** | • *C. difficile* strain detected (GDH+ve); or toxigenic strain detected(NAAT+ve)<br>• No free toxins detected (toxin A/B -ve) | Discordant results – need to assess for the presence of toxigenic strain | Sample needs to be tested with NAAT or TC:<br>• **positive test result** – clinical evaluation is required to distinguish between active infection or asymptomatic carrier of *C. difficile* strain.<br>• **negative result** – CDI is unlikely to be present. |

### 3.5.4 Antibiotic treatment for CDI

Upon confirmation of CDI, antibiotic treatment for CDI should be provided (or continued) based on a combined assessment of: (i) symptoms; (ii) disease severity; and (iii) individual risk factors. Different antibiotic treatment strategies for patients with an initial episode of CDI are available depending on patients' disease severity (181).

DH guidelines emphasise the need to check patient's symptoms daily. If patients remain symptomatic (i.e. with diarrhoea), clinicians should wait until day 7 from the start of treatment regimen before considering this a case of treatment failure (181). For patients with severe CDI and worsening symptoms, DH guidelines recommend a gastroenterology consultation (181).

Where a patient has recurrent CDI (defined as at least 3 consecutive type 5-7 stool within 30 days from previous CDI and a positive *C. difficile* toxin test result) (181), non-*C. difficile* specific antibiotics should be discontinued to allow the normal intestinal flora to recover. In addition, antibiotic treatment for recurrent CDI should be administered taking into consideration symptoms, disease severity, and individual risk factors (223). In the context of multiple recurrences for CDI, clinicians should review all antibiotics and other drug therapies being provided, as well as considering the need to provide additional antibiotic treatment for CDI (181).

### 3.5.5 Infection-control measures

Healthcare staff in contact with patients confirmed and suspected with CDI should always use disposable gloves and aprons, while also following hand-washing policy before and after each patient contact, as per the SIGHT protocol (178).

Ideally, all patients suspected with CDI should be moved into single room isolation as soon as possible. Where there is a shortage of available single rooms, patients confirmed with CDI could be grouped together into a *cohort bay* – an area within an hospital ward where patients confirmed with infectious diarrhoea are separated from other patients using a solid partition and door (178). Alternatively, patients with CDI should be moved into a dedicated *C. difficile* isolation ward with a minimised risk of infection transmission.

If a patient was not initially isolated into a single room upon suspicion, they should immediately be moved into a single room isolation upon receipt of a positive

diagnosis of CDI. This helps to minimise the risk of wide-spread transmission of *C. difficile* pathogen within the hospital ward (178). Patients with confirmed CDI should be released from single room isolation after 48 hours from resolution of diarrhoea (type 1-4 Bristol Stool Chart) (178).

## 3.6 Chapter summary

- This chapter described the clinical context underpinning the thesis case study: new rapid diagnostic tests for CDI.
- Key challenges in the diagnosis were highlighted, including: (i) uncertainty regarding who to suspect and test for CDI; (ii) the number of analytes associated with CDI (e.g. *C. difficile* pathogen, toxins, toxigenic strains); and (iii) a lack of agreement regarding the diagnostic reference test – be it either CCNA or TC.
- The recommended approach to CDI diagnosis in the UK is to run a two-stage testing algorithm for reporting cases of CDI to overcome the shortcomings of each standalone test.
- Upon suspicion of CDI, healthcare staff should put in place infection control measures ranging from hand-washing policy to single room isolation. Antibiotic treatment is provided to patients confirmed with CDI based on a combined assessment of symptoms, disease severity and individual risk factors.

Based on the limitations of current diagnostic tests for CDI, it is crucial that the R&D efforts should be appropriately focused on key areas of unmet clinical need. Chapter 4 outlines the development and results of an online survey of UK healthcare professionals which aimed to identify areas of unmet clinical need for new CDI diagnostics.

# Chapter 4

# Survey to identify clinical needs for new diagnostic tests for *Clostrioidies difficile* infection

## 4.1 Chapter outline

In Chapter 2, the current methodology for developing TPPs for medical tests was presented. Core to the TPP development process is scoping the clinical problem to address, while also mapping the care pathway where new and existing tests might sit. To guide the development of a future TPP for new diagnostic tests for CDI, an online survey of UK healthcare professionals was conducted to identify unmet clinical needs for new diagnostic tests for CDI. This chapter describes the development and dissemination of the online survey, followed by an overview and discussion of the survey results.

This chapter is structured as follows. The aim and objectives of the survey are presented in section 4.2. The structure and development of the survey are outlined in sections 4.3.1 and 4.3.2, followed by a description of the recruitment process (see section 4.3.3), data analysis (section 4.3.4), and ethical approval (section 4.3.5). Results from the survey are presented (section 4.4), and then discussed (section 4.5).

## 4.2 Aim and objectives

An online survey of UK healthcare professionals was conducted to assess if there is an unmet clinical need for new diagnostic tests for CDI in the UK, to inform future R&D activities in this area. The following objectives were addressed:

1. To assess the areas of current unmet clinical need;
2. To map the clinical care pathway for CDI diagnostics; and
3. To identify key characteristics that new tests for CDI should ideally possess (e.g. type of sample, ideal positioning into the clinical pathway, acceptable cost of a test).

## 4.3 Methods

An anonymised online survey was developed in accordance to best practice methods from the Checklist for Reporting Results of Internet E-Surveys (CHERRIES) (238). The sections below outline the structure (see section 4.3.1), validation process (section 4.3.2), dissemination (section 4.3.3), and analysis of the online survey (section 4.3.4). Details on the ethical approval sought for this research are provided in section 4.3.5.

### 4.3.1 Structure

An online survey comprising of four sections was developed in Online Surveys (formerly Bristol Online Survey), including: (1) description of clinical and laboratory practice; (2) problems with diagnosing CDI; (3) room for improvement in current diagnostic tests; and (4) a description of an ideal diagnostic test. Figure 4-1 shows a simplified schematic of the survey structure.

The survey was adapted from the EFLM TE-WG checklist to identify unmet clinical needs for new biomarkers (18). Checklist questions were customised to the disease area of interest where necessary. This checklist presents several disease-agnostic questions which, during the pre-testing session of the survey, were considered difficult to answer by potential respondents without any contextual questions on the disease of interest. The order of survey questions was then tailored to respondents' field of expertise (see Table 4-1 for respondents' subgroups), care setting and knowledge of the disease area using conditional branching (or *skip logic*). This feature of online surveys allows changes to which question(s) a respondent answers next based on their response to a current or previous opening question (239). For example, respondents who stated that they did not think there were any problems with current diagnostic tests for CDI available in their clinical setting would not be asked to list specific limitations with diagnostic tests. A copy of the survey is available in Appendix F.

The following paragraphs describe each of the survey sections. It should be noted that sections 2, 3 and 4 are common to each of the respondent subgroups.

#### 4.3.1.1 Section 1 – Description of Clinical or Laboratory Practice

This section aimed to map the clinical care pathway and testing workflow in place at the care setting where respondents worked. Respondents were asked to map the care pathway via a set of questions tailored to their field of expertise. Respondents were clustered into three subgroups based on their job title and care setting of interest – ward-based clinicians, laboratory-based clinicians and general practitioners (GPs). Clinicians working at the point-of-care and requesting stool testing upon suspicion of CDI were defined as 'ward-based clinicians'; clinicians working in the laboratory and having a greater understanding of diagnostic procedures were labelled as 'laboratory-based clinicians'. A separate subgroup was created for GPs as they worked in a different clinical setting from the other subgroups and might not be familiar with diagnostic tests being requested for CDI – as per expert opinion. Table 4-1 shows the three subgroups of respondents.

**Table 4-1 Respondents' subgroups based on job title and care setting where they work**

| Ward-based clinicians | Laboratory-based clinicians | General Practitioners |
|---|---|---|
| Consultant | Medical Microbiologist | Hospital-based GPs |
| Trainee/Staff Doctor | Biomedical Scientist | Community-based GPs |
| Infectious Disease Doctor | Healthcare Scientist | |
| Nurse | Clinical Scientist | |
| Infection Control Nurse | | |

This survey section was divided into two subcomponents: (1) description of clinical practice; and (2) description of diagnostic pathways. Each of these components is described below, separately.

*Description of clinical practice*

Ward-based clinicians and GPs were asked to describe: (i) which symptoms would lead them to suspect CDI; (ii) the age group in which they would most encounter individuals suspected with CDI; and (iii) the implementation and de-

escalation of any infection-control measures in place for patients suspected and confirmed with CDI.

*Description of diagnostic pathways*

Ward- and laboratory-based clinicians were asked to describe which diagnostic tests were available in their clinical setting, either standalone tests or testing algorithm. Laboratory-based clinicians were invited to explain why a certain testing algorithm was adopted, whereas ward-based clinicians were requested to rate their confidence in test results for CDI.

#### 4.3.1.2 Section 2 – Problems with Diagnosing CDI

Each respondent was asked if there were any problems affecting the diagnosis of CDI and then, specifically, if there were any issues with CDI diagnostics available in their clinical setting and what problems these limitations might cause.

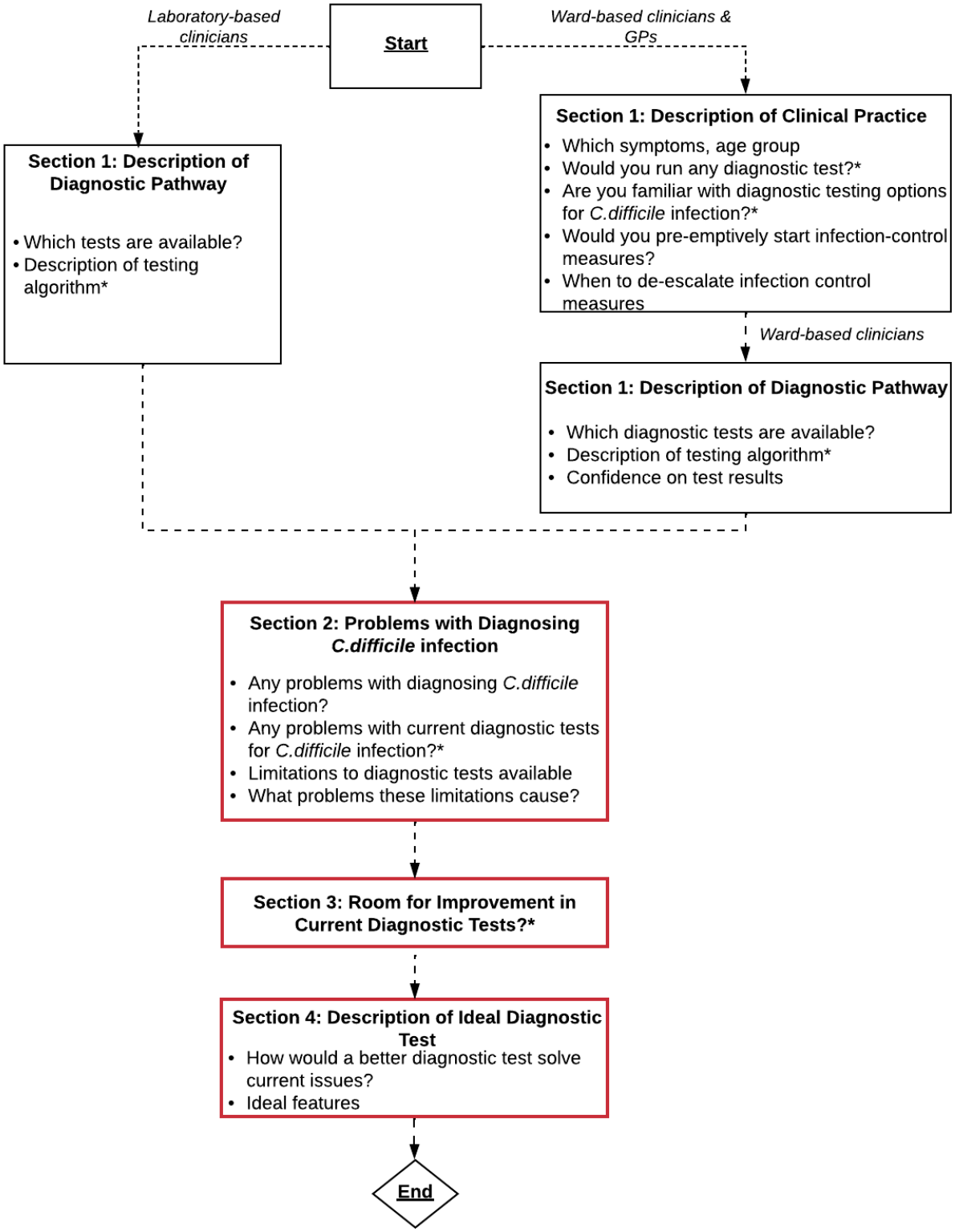
#### 4.3.1.3 Section 3 – Room for Improvement in Current Diagnostic Tests?

This section aimed to assess if respondents believed there was scope for improvement in current diagnostic tests for CDI. If respondents answered ‘yes’, they progressed to Section 4 or, if they answered ‘no’, they were asked to explain their previous answer and then ended the survey. As some respondents might regard current diagnostic tests for CDI as satisfactory, it was important to understand the rationale underlying their viewpoint.

#### 4.3.1.4 Section 4 – Description of Ideal Diagnostic Test

This section was designed to gauge respondents’ views on the ideal characteristics for a new diagnostic test for CDI, including: (i) type of sample; (ii) technique for obtaining the sample; (iii) method for transporting the sample; (iv) turnaround time of testing; (v) positioning of the new test into the existing care pathway; and (vi) test cost. It was decided against seeking respondents’ views on the ideal diagnostic accuracy for tests, due to an expected variable understanding of test accuracy among healthcare professionals (153).

**Figure 4-1 Schematic of the online survey structure. Sections in red are common to all respondent subgroups (e.g. ward-based clinicians, laboratory-based clinicians and GPs). Sections with asterisk (*) indicate conditional branching questions.**



### 4.3.2 Validation

The survey was validated through an iterative pre-testing and piloting process (240-242), including:

1. **informal pre-testing** – potential respondents were asked to indicate on a feedback form whether each question was sufficiently understandable, without filling in the survey (n=7). The survey wording was edited after each round of pre-testing.

2. **expert-driven pre-testing** – consultations were then conducted with clinical staff from the Healthcare Associated Infection Research team at the Leeds Teaching Hospitals NHS Trust (LTHT) (n=3), to ensure that survey wording reflected the appropriate terminology for the disease area of interest. The clinical experts were asked to *think-aloud* while reviewing the survey without completing any feedback form. As above, the survey wording and structure was edited after each round of pre-testing to receive as much feedback as possible.

3. **piloting** – the survey was piloted between February 2020 and June 2020 with a different group of clinical experts (n=5), to ensure usability and technical functionality of the survey in terms of access, navigations and submission.

### 4.3.3 Dissemination

The survey was targeted at medical doctors, nurses, healthcare scientists and GPs with experience of CDI, across hospital and community care settings in the UK.

The survey was made available during two rounds of dissemination: (i) between 21st June 2020 and 19th November 2020; and (ii) between 29th April 2021 and 1st June 2021. Two rounds of dissemination were conducted to account for the significant pressure the targeted respondents, namely healthcare professionals who work in infection prevention and control, and laboratories, were facing during the COVID-19 pandemic. In November 2020, another COVID-19 wave was underway and the NHS was under significant pressure, it was therefore decided to pause dissemination and relaunch the survey in April 2021. The survey was disseminated online via: (1) social media platforms (Twitter and Facebook); and (2) three separate mailing lists of UK clinical contacts, National Institute for Health

and Care Research (NIHR) Leeds In Vitro Diagnostic Co-operative (243), RID-AMR@Leeds research group (89) and Combatting Bacterial Resistance in Europe- CDI (COMBACTE-CDI) study (244).

The survey was open to anyone who voluntarily clicked the survey link with no incentives being offered to respondents. The survey took approximately 5-10 minutes to complete and did not require respondents to search for any additional information. Respondents were asked to answer several mandatory questions which were highlighted by an asterisk – these questions always included a non-response option such as 'I do not know'. Respondents were able to change their answers via a 'Back' button.

As the survey was initially set for dissemination in April 2020, the dissemination strategy was significantly revised in response to the COVID-19 pandemic. The initial dissemination strategy was based on snowball sampling to identify respondents that fitted key selection criteria (245, 246). These respondents, in turn, were asked to send the survey link via email to their professional contacts until data saturation (245, 246). Between January and March 2020, several UK-based thought leaders[14] in the field of CDI diagnosis were contacted via phone call to ask them to share the survey to their colleagues (n=9). While some of them initially agreed to help, in April 2020 most contacts were unable to disseminate the survey due to the severe pressures clinical staff were facing in light of the COVID-19 pandemic. It was therefore decided to simplify the recruitment strategy, focusing on dissemination via existing clinical networks and social media, while also postponing the launch of the survey.

[14] In January 2020 Professor Mark Wilcox, one of the clinical investigators from the Leeds RID-AMR research team, provided the candidate with a list of UK-based thought leaders in the field of CDI diagnosis.

### 4.3.4 Data analysis

Survey data was analysed and sorted by respondent groups (ward- and laboratory-based clinicians, and GPs). Descriptive statistics (e.g. percentages and frequencies) were calculated for close-ended questions, whilst thematic analysis was conducted for open-ended questions. Themes embedded within respondents' answers were drawn directly from the data, or from the key evidence domains underpinning the test evaluation pathway – see Table 4-2. These evidence domains were applied to cluster test characteristics included within TPPs into common themes – as presented in Chapter 2.

Data processing and analysis were conducted in Microsoft Excel – with only completed questions being included in the analysis. A single analyst (Cocco P) conducted all data analysis, with any uncertainties discussed with the additional researchers (Davies KA, Shinkins B, Smith AF, West RM) as required.

**Table 4-2 Definition of test evidence domains applied during thematic analysis.**

| Test evidence domain | Definition | Source |
|---|---|---|
| **Target test population** | Factors related to the disease of interest, and selection of which individuals should be tested. | (23, 247) |
| **Pre-analytical factors** | Factors related to the "patient preparation, sample collection, handling, transportation, storage and preparation of testing". | (247) |
| **Analytical performance** | The performance of a test concerning the analytical phase (i.e. the point of sample analysis). It is the ability of a test to correctly detect and measure a particular analyte (e.g. analytical precision, limits of detection). | (27, 94) |
| **Clinical validity** | Ability of a device to yield results that are correlated with a particular clinical condition or a physiological or pathological process or state. | (27) |
| **Clinical utility** | Ability of a test to affect relevant health-related outcomes for patients (e.g. improvement in quality of life, longer lifespan). | (30) |
| **Post-analytical factors** | Factors related to the correct interpretation and reporting of test results. | (95, 248) |

### 4.3.5 Ethical approval

As the survey collected views from NHS staff who were recruited in virtue of their professional role, advice was sought from the School of Medicine Research Ethics Committee (SoMREC) at the University of Leeds as to whether NHS Health Research Authority (HRA) ethical approval was needed (see Appendix G for more information). The ethical approval from SoMREC was deemed appropriate since this study was not considered 'research' according to the NHS as: (i) it did not involve participant randomisation nor change in clinical care; and (ii) survey findings were not generalisable to other health conditions apart from CDI. This survey was then approved by the SoMREC at the University of Leeds (application reference number MREC 19-047 – see Appendix H).

The survey was fully anonymised with no direct identifiers of participants (e.g. name, contact details) being asked. Data contained two types of indirect identifiers – job title and country where respondents worked – which did not enable to identification of respondents. Data collected from this survey was not classed as sensitive or controversial.

Respondents were asked to give their consent to participate to the study after reading an information sheet provided on the first page of the survey. The information sheet contained details regarding: (i) the purpose of the survey; (ii) the length of time required to complete the survey; (iii) which data was stored; and (iv) who had access to the data for conducting the analysis. After completing the survey, participants were asked again to confirm their consent to submit their answers; participants were also informed that upon submission their responses could not be withdrawn. Data was safely stored in the University of Leeds One drive.

As the dissemination date in the ethics application for the survey was originally set to September 2020, two amendments to the ethics application were submitted to extend the dissemination period until May 2021 and, subsequently, until August 2021, to maximise the number of responses.

## 4.4 Results

Additional results of the survey results are provided in Appendix I. Key findings are summarised herein. Note that, as certain survey sections entailed conditional branching, the number of participants answering each question may vary.

### 4.4.1 Participants

The online survey was disseminated from 21st June 2020 until 1st June 2021 receiving a total of 48 responses. Table 4-3 summarises the survey participants. Most of the respondents were infection control nurses (n=21, 44%), followed by consultants (n=9, 19%), and healthcare scientists (n=8, 16%). The most common clinical setting was the hospital (n=38, 79%), with most respondents working in England (n=42, 88%). Few respondents reported that they worked in local health authorities (n=3, 6.3%) or Public Health – Health Protection (n=2, 4.2%).

**Table 4-3 Characteristics of survey participants (n=44) in absolute values (n) and as percentages (%)**

| **Job title** | **n (%)** |
|---|---|
| Infection control Nurse | 21 (44) |
| Consultant | 9 (19) |
| Healthcare Scientist | 8 (16) |
| General Practitioner | 4 (8) |
| Infectious Disease Doctor | 2 (4) |
| Nurse | 2 (4) |
| Medical Microbiologist | 1 (2) |
| Other | 1 (2) |
| **Subgroups** | |
| Ward-based clinician | 33 (69) |
| Laboratory-based clinician | 11 (23) |
| GP | 4 (8) |
| **Years of work experience** | |
| 0 to 4 years | 8 (17) |
| 5 to 10 years | 6 (13) |
| 11 to 15 years | 10 (21) |
| More than 15 years | 24 (50) |
| **Country** | |
| England | 42 (88) |
| Scotland | 3 (6) |
| Northern Ireland | 1 (2) |
| Wales | 2 (4) |
| **Clinical setting** | |
| Hospital | 38 (79) |
| GP medical practice | 5 (10) |
| Other | 5 (10) |

### 4.4.2 Description of clinical practice

All ward-based clinicians (n=33) and GPs (n=4) were asked the symptoms that would raise suspicion for CDI and which patient age groups were typically suspected of having CDI in their clinical setting.

Figure 4-2 shows the symptoms that respondents stated would raise suspicion of possible CDI. Out of 37 respondents, watery stools (95%, n=35) and unexplained diarrhoea (n=35, 95%) were the most common symptoms leading to suspicion of CDI across each respondent subgroup. Among ward-based clinicians and GPs (n=37), older adult patients were reported as the most common patient age group suspected of CDI (n=34, 92%), followed by patients aged between 18 to 65 years (n=8, 22%) – as presented in Figure 4-3. Younger patients were rarely suspected of having CDI (n=3, 8%).

Except for two respondents, all ward-based clinicians and GPs (n=35) stated that they would request a diagnostic test upon suspicion of CDI and would pre-emptively start infection-control measures while waiting for test results – such as single room isolation in the hospital setting, or hand hygiene at the community setting.

**Figure 4-2 Symptoms leading to suspicion of CDI in absolute number (n) and as a proportion of respondents selecting that option within each subgroup (%). The subgroup 'Nurses' comprises responses from Nurses (generic) and Infection Control Nurses, whereas the subgroup 'Others' comprises responses from a General Practitioner and a Public Health officer.**

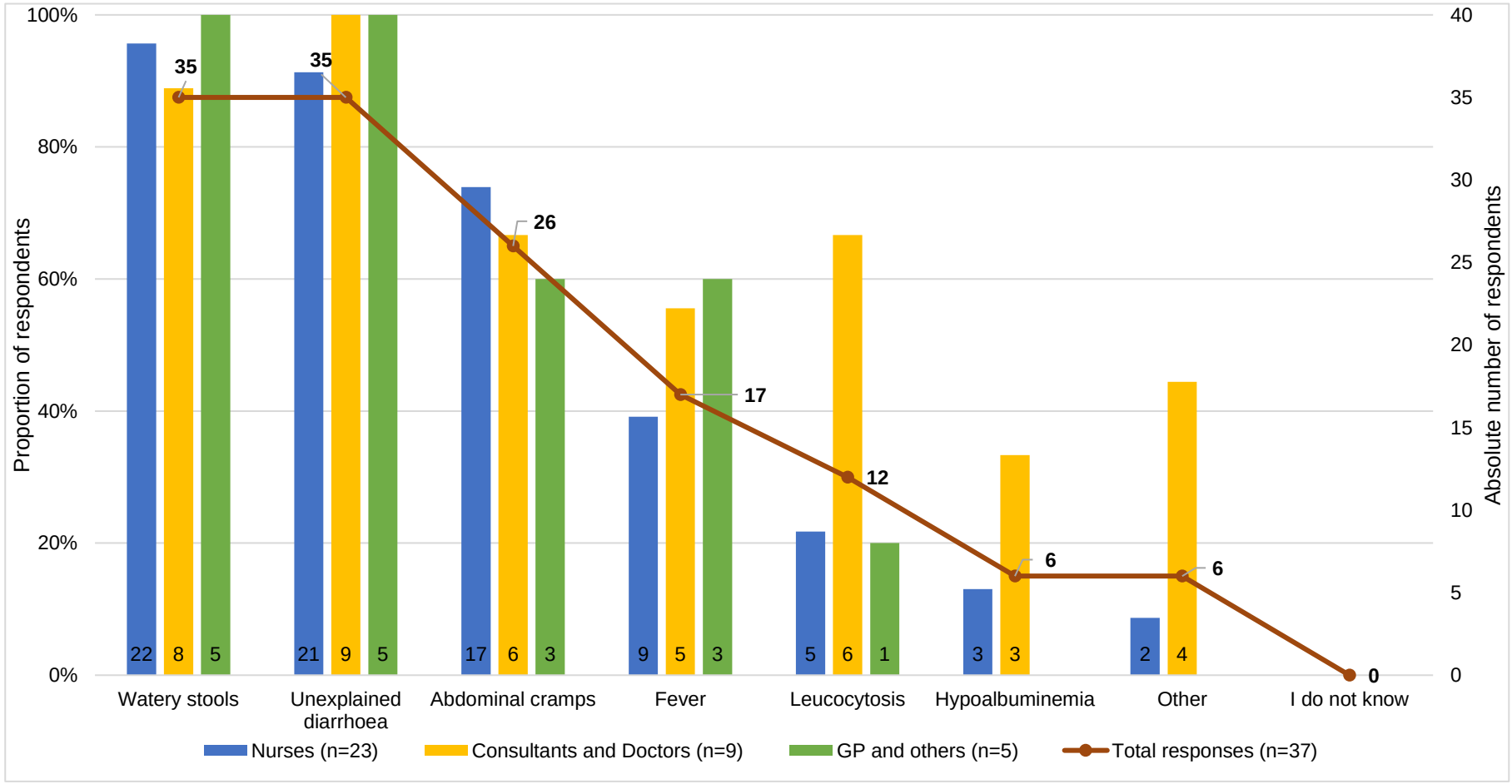
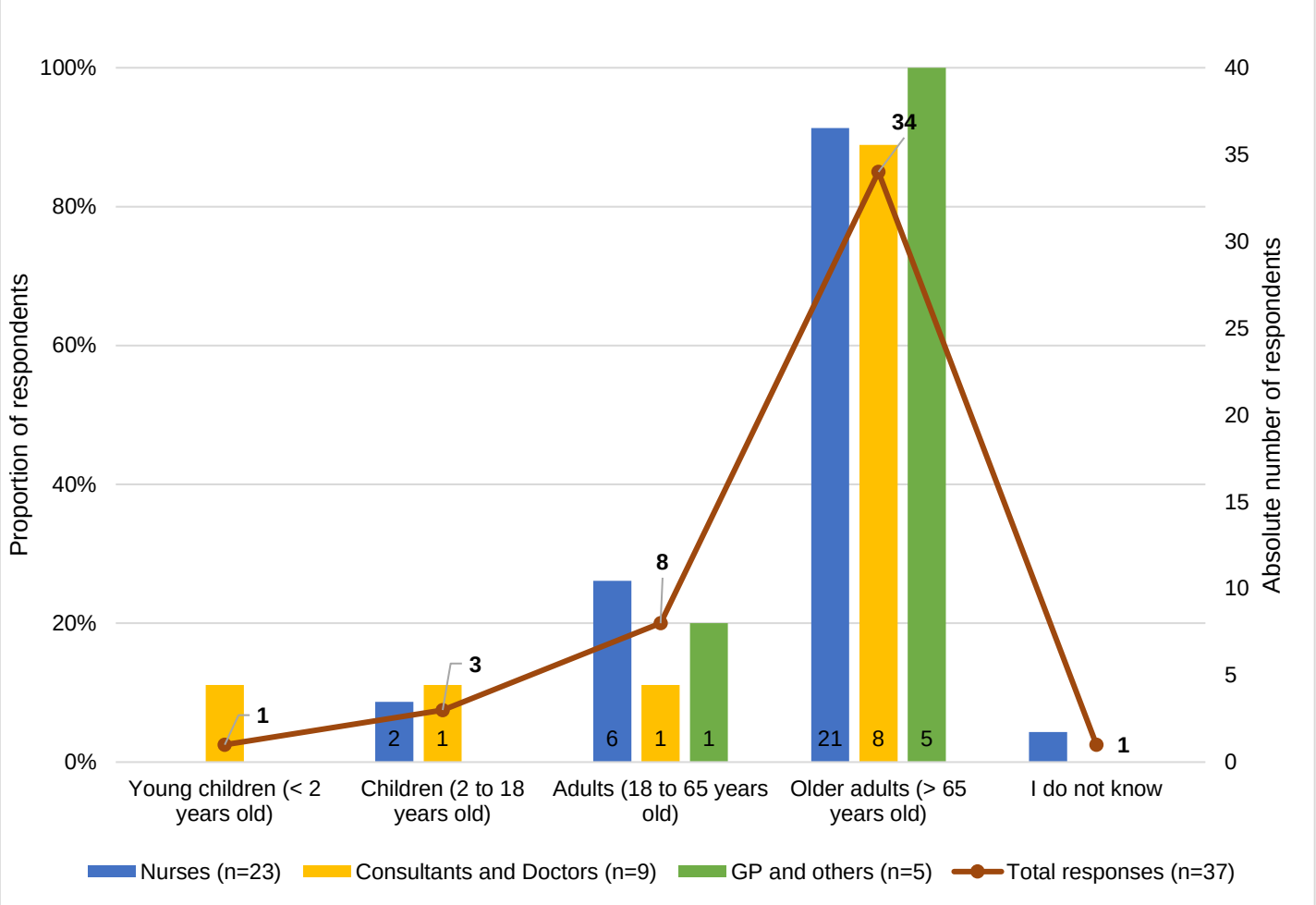


**Figure 4-3 Patient age groups most often suspected for CDI in absolute number (n) and as a proportion of respondents selecting that option within each subgroup (%).**



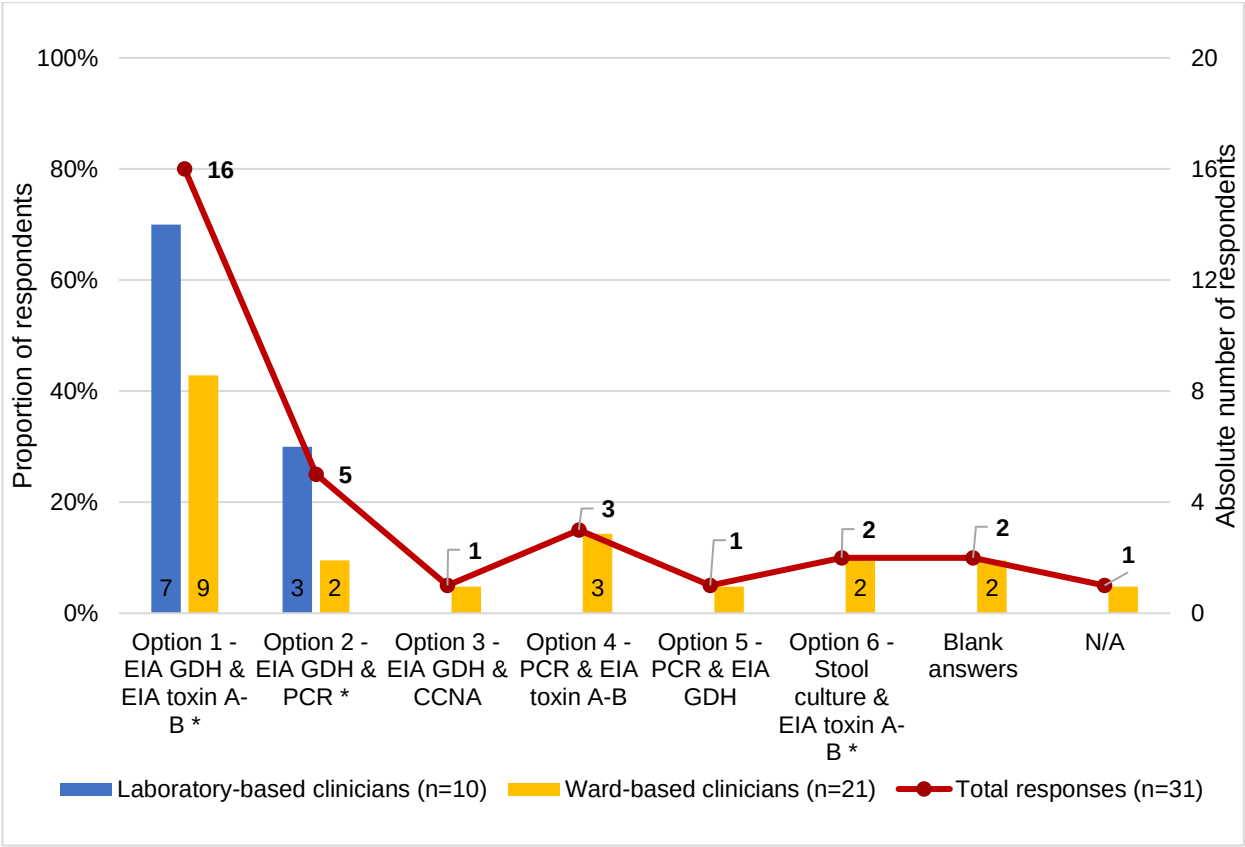

### 4.4.3 Description of diagnostic pathway

Among ward-based clinicians who would request testing for CDI (n=33), the majority reported being familiar with the test(s) the laboratory used to diagnose CDI in their clinical setting (n=29, 88%). The remaining respondents (n=4, 12%) did not answer the next questions on the diagnostic pathway, and moved to the section on Confidence in test results for CDI. Laboratory staff (n=11) and ward-based clinicians familiar with the diagnostic options for CDI (n=29) were asked to describe the laboratory practice for diagnosing CDI in the clinical setting where they worked. Across ward- and laboratory-based clinicians (n=40), EIA for toxin A/B (n= 30, 75%) and GDH (n=29, 73%) were the most common diagnostic tests stated as being available for CDI, followed by PCR (n=19, 48%). Fewer respondents reported the use of culture (n=9, 23%), TC (n=6, 15%) or CCNA (n=3, 8%) to diagnose patients suspected with CDI.

#### 4.4.3.1 Testing algorithm for CDI

Laboratory staff (n=11) and ward-based clinicians familiar with testing options for CDI (n=29) were asked to describe the testing algorithm used in their clinical setting for diagnosing CDI, if any. Among 40 respondents, the majority of ward- and laboratory-based clinicians stated that they would request a testing algorithm for diagnosing CDI (n=31, 78%), with a variety of different testing strategies reported – as shown in Figure 4-4. Across 31 respondents, running EIA GDH and EIA toxin A/B (i.e. 'Option 1') was the most common testing algorithm reported (n=16, 52%), followed by running EIA GDH and PCR (i.e. 'Option 2') (n=5, 16%), and PCR combined with EIA toxin A/B (i.e. 'Option 4') (n=3, 10%). The least common option was stool culture followed up by EIA toxin A/B (n=2, 6%). The following sections describe the different testing algorithms which emerged from responses, including variations within each testing algorithm. It should be noted that variations across testing algorithms are presented based on the first test being run as part of the testing algorithm – either EIA GDH, PCR or stool culture – rather than in order of popularity.

**Figure 4-4 Testing algorithm options for CDI in absolute number (n) and as a proportion of respondents selecting that option within each subgroup (%). Options with asterisk (*) might entail a third confirmatory test.**



**Option 1 – EIA GDH & EIA toxin A/B**

The most common testing algorithm among ward- and laboratory-based clinicians was testing with EIA GDH, then following up with EIA toxin A/B (n=16). Figure 4-5 provides a simplified schematic of testing algorithm 'Option 1'. Some variations in how to run this testing algorithm were found:

- **order of testing options** – EIA GDH and EIA toxin A/B could be run either simultaneously (n=4), or sequentially with EIA toxin A/B being run following a GDH positive test result (n=12); and
- **need for confirmatory testing** – when the pathogen was detected (GDH positive), many respondents would check for the presence of toxigenic strains of the pathogen using PCR for every sample tested with EIA toxin A/B (n=2), or when free toxins were not found in stools (i.e. EIA toxin A/B negative) (n=12). Only two respondents reported no further testing after the first two steps.

This testing algorithm is often used since it is currently recommended by PHE guidance, as three laboratory-based clinicians explained. 'Option 1' testing algorithm appeared to lead to lower testing costs (n=2) – using EIA toxin A/B allows to test more than one sample at once (i.e. *batch testing*) which, in turn, might lead to significant cost savings compared to following-up every sample positive to GDH with PCR. One respondent suggested that this testing algorithm is a clinically- and cost-effective option for diagnosing patients suspected with CDI. Additional advantages of this testing algorithm lied in the reduced time-to-diagnosis and increased ease of use compared to culture.

**Option 2 – EIA GDH & PCR**

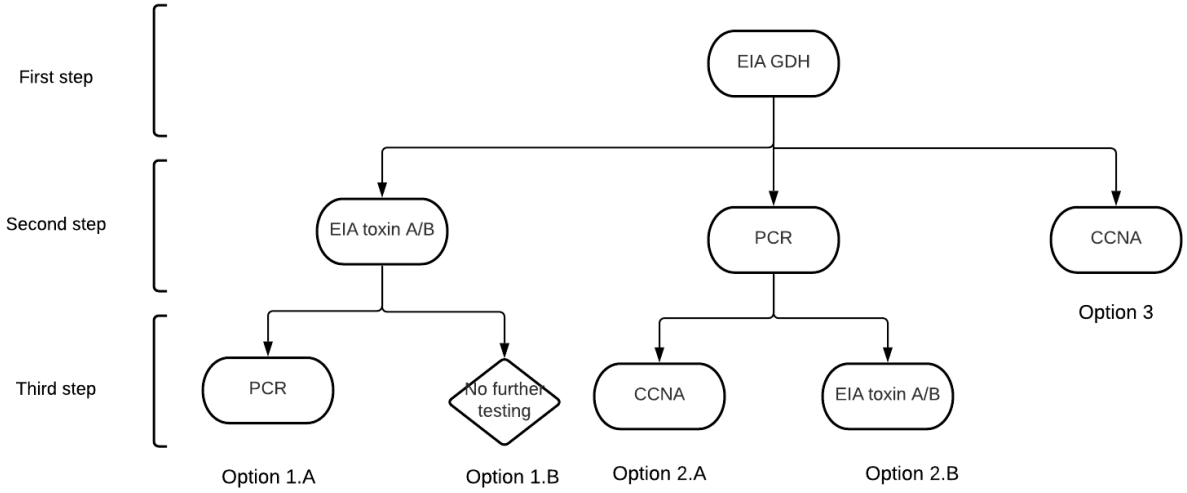
Another common testing algorithm among ward- and laboratory-based clinicians was testing first with EIA GDH, then following-up with PCR (n=5). Figure 4-5 provides a simplified schematic of testing algorithm 'Option 2'. Main differences as to how this testing algorithm was run were based on:

- **order of testing options** – EIA GDH and PCR could be run simultaneously (n=1), or in a sequence (n=3); and
- **choice of test for toxins** – since detecting the presence of the pathogen with GDH or toxigenic strains with PCR is insufficient to confirm the

presence of CDI, it is recommended to run an additional test for toxins – be it either CCNA (n=2) or EIA toxin A/B (n=2).

Running this testing algorithm helps to quickly de-isolate a non-infectious patient who has tested positive to GDH but negative to PCR, as a laboratory-based clinician explained. Laboratories affiliated with the *C. difficile* ribotyping network (CDRN) are required to run CCNA (n=1), since CCNA is currently one of the gold standards for detecting toxins in stools, as two respondents stated. The combination of EIA GDH, PCR and following-up with EIA toxin A/B appeared to accurately identify carriers and infected cases.

**Figure 4-5 Simplified schematic of Option 1, 2 and 3 testing algorithms including of variations within each option**




### Option 3 – EIA GDH & CCNA

One respondent suggested testing first with GDH EIA, and if positive to follow-up with CCNA – see Figure 4-5.

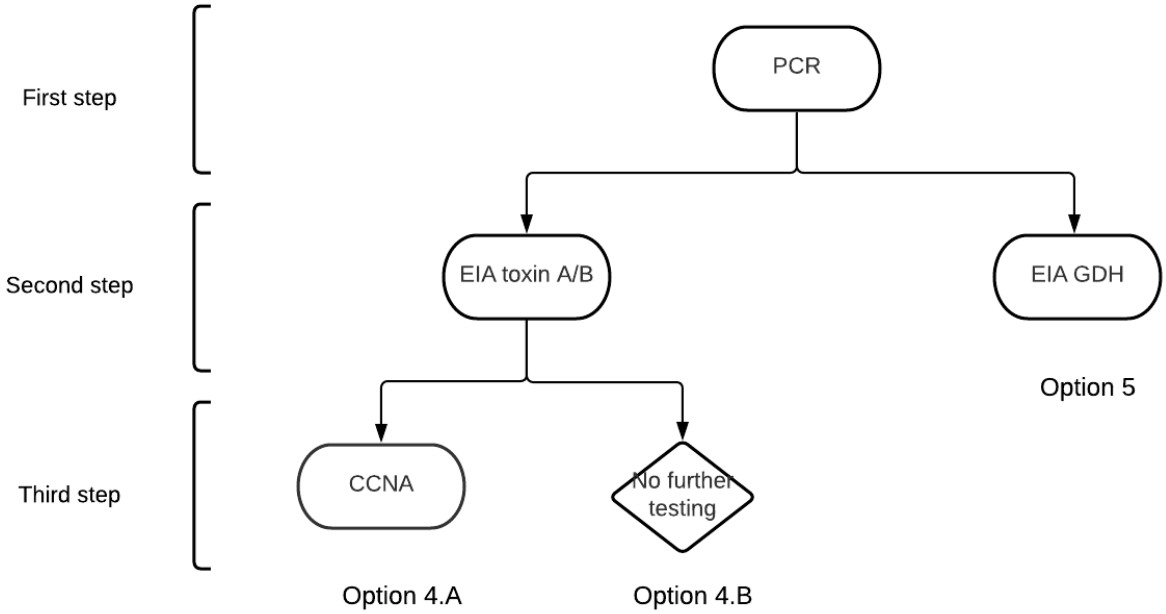
### Option 4 – PCR & EIA toxin A/B

A lesser common testing algorithm option was testing first with PCR, then following-up with EIA toxin A/B (n=3). There was some variation as to whether CCNA is used as a confirmative third testing step (n=2) or not (n=1). Since the reported diagnostic sensitivity of EIA toxin A/B is low, it is recommended to run a confirmatory test for toxins to increase confidence in the combined test result. Figure 4-6 illustrates a simplified schematic of the 'Option 4' testing algorithm.

### Option 5

One respondent reported testing first with PCR and then EIA GDH (see Figure 4-6).

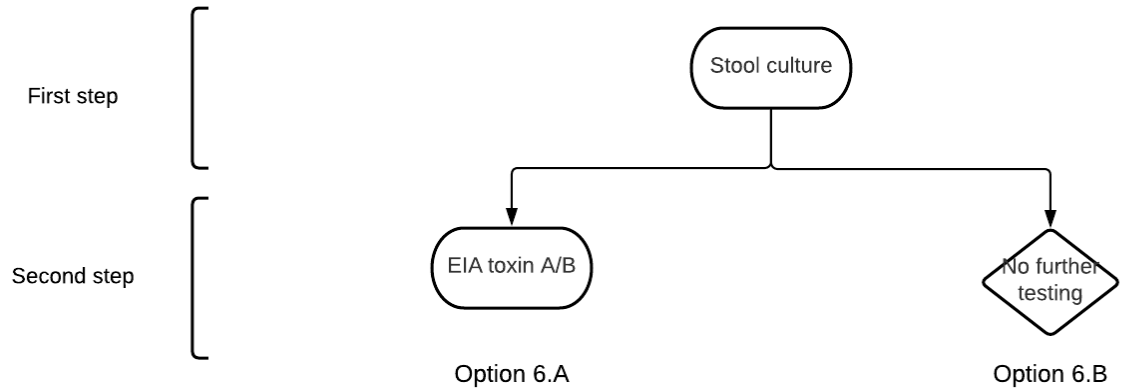
**Figure 4-6 Simplified schematic of Option 4 and 5 testing algorithm including variations within each testing algorithm option**




### Option 6

An additional testing algorithm option was to test first with stool culture and then either follow-up with EIA toxin A/B (n=1) or not (n=1). Figure 4-7 gives a simplified schematic of 'Option 6' testing algorithm.

**Figure 4-7 Simplified schematic of Option 6 testing algorithm including variation within this testing algorithm option**

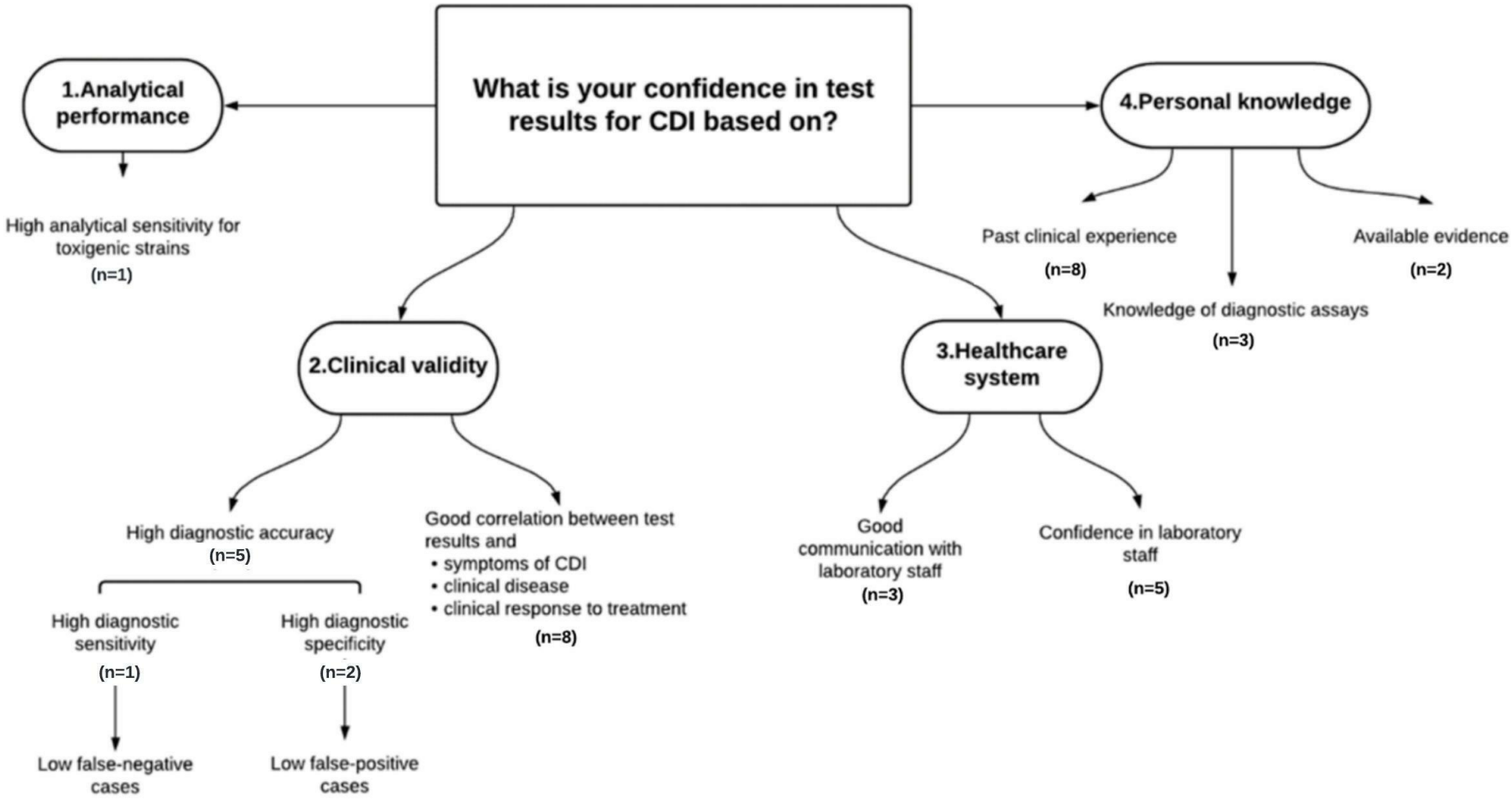


#### 4.4.3.2 Confidence in test results for CDI

All ward-based clinicians (n=33) and GPs (n=4) were asked to state how confident they usually feel about the test results for CDI they receive. Among 37 respondents, the majority reported that they were ‘Fairly confident’ (n=17, 46%) or ‘Very confident’ (n=17, 46%) about the test results for CDI they receive.

All ward-based clinicians (n=33) and GPs (n=4) were asked to state what their confidence in test results for CDI was based on. Twenty-eight respondents answered this question. Four major themes emerged from this qualitative question (see Figure 4-8) – two themes were based on the key evidence domains underpinning the test evaluation methodology, whereas the others were directly drawn from the data. The following sections describe each of the themes identified.

**Figure 4-8 Key factors giving ward-based clinicians and GPs confidence in test results for CDI, sorted by common themes**



**Theme 1 – Analytical performance**

One respondent mentioned that their confidence in test results for CDI was based on the high analytical sensitivity for detecting toxigenic strains in patients' stool samples.

"Test very sensitive, only detects toxigenic strains of C.diff [sic]" – *Infection Control Nurse working in an English hospital*

**Theme 2 – Clinical validity**

Eight respondents stated that their confidence in testing results for CDI was based on high diagnostic accuracy of the testing options being used – specifically a high diagnostic sensitivity which, in turn, leads to a low incidence of missed cases (i.e. false-negative patients [FN]), and a high diagnostic specificity which leads to few FP patients.

"Haven't [sic] known any false positives" – *Infection Control Nurse working in an England-based local authority*

"based on clinical responses and few cases being missed i.e. no large numbers of clinical deterioration due to false negatives" – *Consultant in Infectious Diseases working in an English hospital*

Correlation between test results and CDI clinical disease was often perceived as a strong indicator of the test results for CDI being correct (n=8) – especially in the context of positive test results. Specifically, a good correlation between test results and (i) clinical symptoms; (ii) clinical disease of CDI; or (iii) clinical response to CDI treatment, made the respondents more confident in test results for CDI.

"Positive results in combination with clinical signs are strongly indicative of CDI, negative results are less reliable" – *Infection Control Nurse working in a Welsh hospital*

"Clinical Symptoms [sic] match up with lab diagnosis" – *Infection Control Nurse working in an England-based Community Mental Health Trust*

**Theme 3 – Healthcare system**

Three respondents explained that their confidence in test results was based on good communication between laboratory staff and ward-based clinicians, alongside confidence and support in the laboratory staff and their guidance (n=5).

"confident with the system in use and the good communication from medical staff ad lab staff" – *Infection Control Nurse working in an English hospital*

"[..] high quality lab support within my organisation" – *Consultant in Infectious Diseases working in an English hospital*

**Theme 4 – Personal knowledge**

For eight respondents, confidence in test results for CDI came from their past clinical experience and knowledge of diagnostic tests. Two respondents also referred to the published evidence and rigorous research underpinning the NHS testing practices.

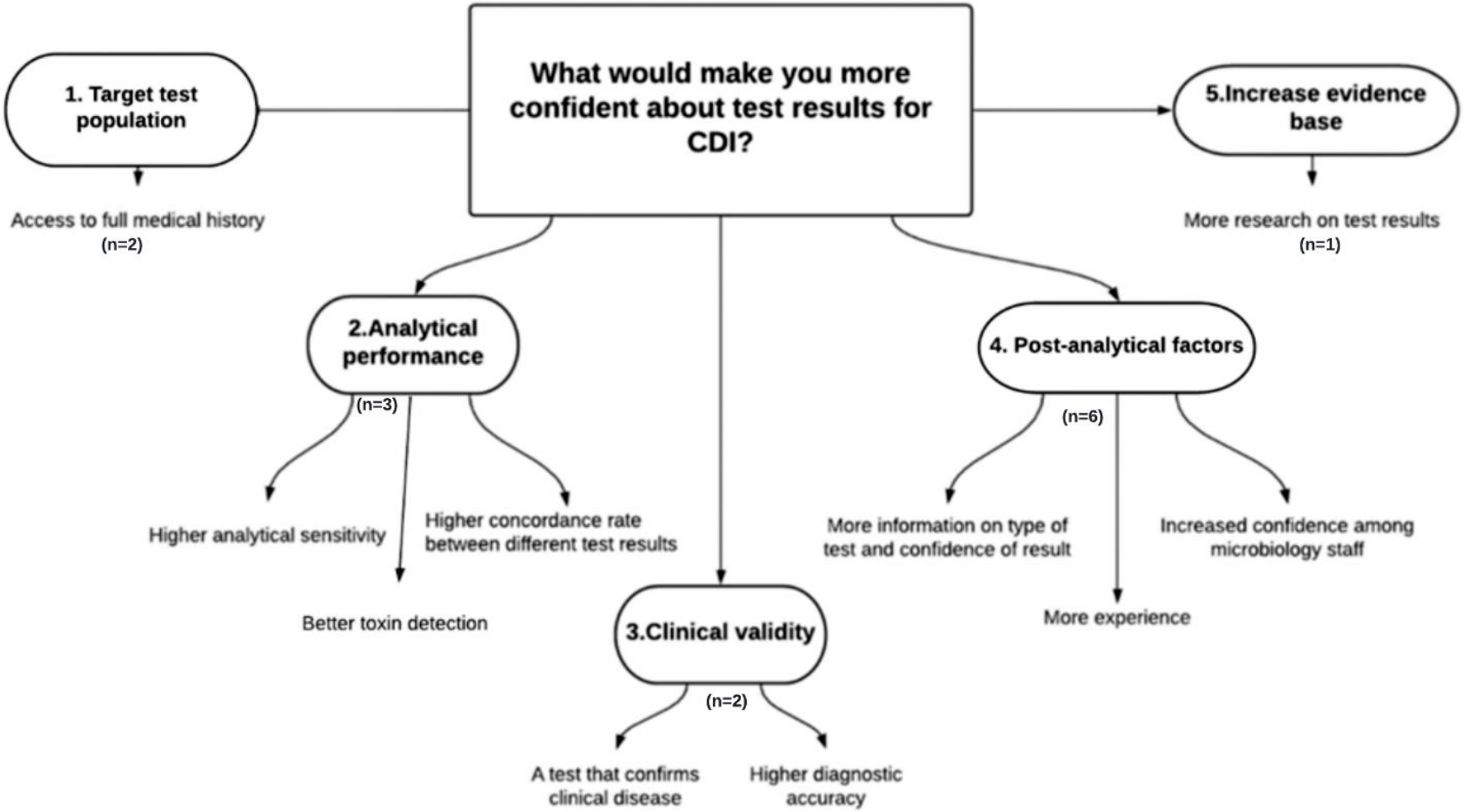
"Knowledge of the assays and their characteristics, combined with clinical experience" – *Infection Disease Doctor working in an English hospital*

"Diagnostic testing used in the NHS is usually evidence based, following rigorous research" – *Public Health – Health Protection working in England*

#### 4.4.3.3 Factors increasing confidence in testing results

All ward-based clinicians (n=33) and GPs (n=4) were asked to state what would make them more confident about test results for CDI. Among the respondents who answered this question (n=22), five themes emerged from the data and the key evidence domains underpinning test evaluation methodology. Out of 37 respondents, many respondents stated 'Do not know' or 'Nothing' or did not answer (n=15, 41%). Figure 4-9 shows the main factors that would increase confidence in test results for CDI among ward-based clinicians and GPs, sorted by common themes.

**Figure 4-9 Key factors that would make ward-based clinicians and GPs more confident in test results for CDI, sorted by common themes**



**Theme 1 – Target test population**

Two respondents stated that access to patients' full medical history – such as medications being prescribed to patients – might increase their confidence in test results. As diarrhoea is a common side effect caused by many treatments (e.g. antibiotics), clinicians need full access to patients' medical history to rule out with confidence an episode of diarrhoea due to non-infectious causes. Another respondent suggested an increased ability to interpret clinical symptoms would increase their confidence in test results for CDI.

"Full medical history including medications prescribed" – *Infection Control Nurse working in an English hospital*

**Theme 2 – Analytical performance**

Two respondents mentioned that better detection of toxins in patients' stools would boost their confidence in test results for CDI, whereas another highlighted the need for an increased concordance rate between tests run as part of the same testing algorithm.

"Better detection of toxin" – *Infection Control Nurse working community setting in Wales*

"Fewer equivocal results" – *Infection Control Nurse working in a Scottish hospital*

**Theme 3 – Clinical validity**

One respondent suggested that a non-invasive test confirming likely clinical disease of CDI might help to increase their confidence in test results for CDI.

"A test to confirm likely clinical disease that is not invasive – sigmoidoscopy could be used to confirm but is too invasive. PCR for toxin gene may help but needs to be tied to clinical features and may overdiagnose and may lead to misdiagnosis" – *Consultant in Infectious Diseases working in an English hospital*

Another respondent proposed a standalone test with high diagnostic accuracy and quick turnaround time, coupled with good correlation with clinical disease.

"A single highly sensitive and specific test that is quick. A test that correlated with clinical picture" – *Consultant in Infectious Diseases working in an English hospital*

**Theme 4 – Post-analytical factors**

Five respondents suggested that more information on the type of diagnostic test being requested, and confidence intervals of test results might enhance their trust in test results for CDI, or alternatively more experience and knowledge in diagnosing patients with CDI. Another respondent also suggested an increase in confidence among microbiology staff.

"Some understanding of the tests carried out to diagnose C.diff" – *Infection Control Nurse working in an English hospital*

"Additional info – type of test used and confidence of result" – *Public Health – Health Protection working in England*

**Theme 5 – Increased evidence base**

One respondent suggested that more evidence-based research on test results of CDI diagnostics might help to strengthen their trust in test results for CDI.

#### 4.4.3.4 De-escalation of infection control measures

Ward-based clinicians were asked when they would start de-escalating infection control measures for a patient confirmed with CDI (n=33). The majority of the respondents would make this decision based mainly on symptoms (n=25, 76%), followed by treatment options provided to patients (n=6, 18%). Several respondents indicated they would take into account other system-related aspects (n=5, 15%) such as availability of single rooms or completion of deep cleaning of a given isolation room. Three ward-based clinicians stated they would also consider the full clinical picture of patients positive to CDI (9%), including likelihood of disease relapse, presence of other underlying or concurrent conditions or absence of colitis.

For those ward-based clinicians who reported symptoms as the main factor which would prompt removal of infection-control measures (n=25), 48 hours from absence of loose stools and resolution of formed stools was often referred to as the key factor leading to release of patients from single room isolation (n=19, 76%) and generic symptoms (n=4, 16%). One respondent reported that patients positive to CDI would not be de-isolated until they had 7 days of formed stools.

### 4.4.4 Problems with diagnosing *C. difficile* infection

#### 4.4.4.1 Problems with diagnosing CDI

Each of the total survey participant (n=48) were asked, in an open-ended question, to list any issues currently affecting the diagnosis of CDI – excluding those problems relating to diagnostic tests for CDI (discussed separately in section 4.4.4.2). The following sections, together with Figure 4-10, describe the main themes and subthemes that emerged from the data.

**Theme 1 – Target test population**

Seventeen respondents commented on the uncertainty surrounding who and when to test for CDI due to a lack of reliable clinical recognition of CDI. Possible explanations provided for this included:

- **difficulties in distinguishing diarrhoea due to CDI from non-infectious diarrhoea** – diarrhoea is a symptom common to many health conditions and treatments – such as antibiotics – thereby making it challenging for clinicians to rule in CDI with confidence, based on episodes of diarrhoea alone;
- **poor access to patients' full medical history** – many respondents stressed the challenges in getting a complete clinical background on patients. Clinicians often rely on patients' self-reporting their current medications, which may lead to incorrect or incomplete information on which antibiotics (if any) patients are receiving; and
- **over-reliance on stools frequency and volume** – two respondents commented on the excessive reliance on stool frequency to suspect CDI, rather than considering patients' underlying conditions and treatments. As clinicians might have an incomplete understanding of patients' clinical history, suspicion of CDI is often based on the frequency and volumes of stools. Relying on stool frequency, however, was perceived as an inconsistent threshold for suspecting CDI – especially in the context of patients being unwilling to report an episode of diarrhoea promptly.

**Theme 2 – Pre-analytical factors**

Most of the respondents referred to pre-analytical testing factors as key issues hindering the diagnosis of CDI (n=19). Eleven respondents highlighted difficulties in obtaining adequate stool samples from patients – especially those suffering from incontinence. Six respondents also mentioned significant delays in sending stool samples to the laboratory for testing.

**Theme 3 – Analytical performance**

Five respondents regarded aspects relating to the analytical performance of tests as an issue affecting the diagnosis of CDI – such as delayed time-to-yield test result (n=4), or degradations of toxins over time (n=1). The latter can reduce the analytical sensitivity for detecting toxins in stools.

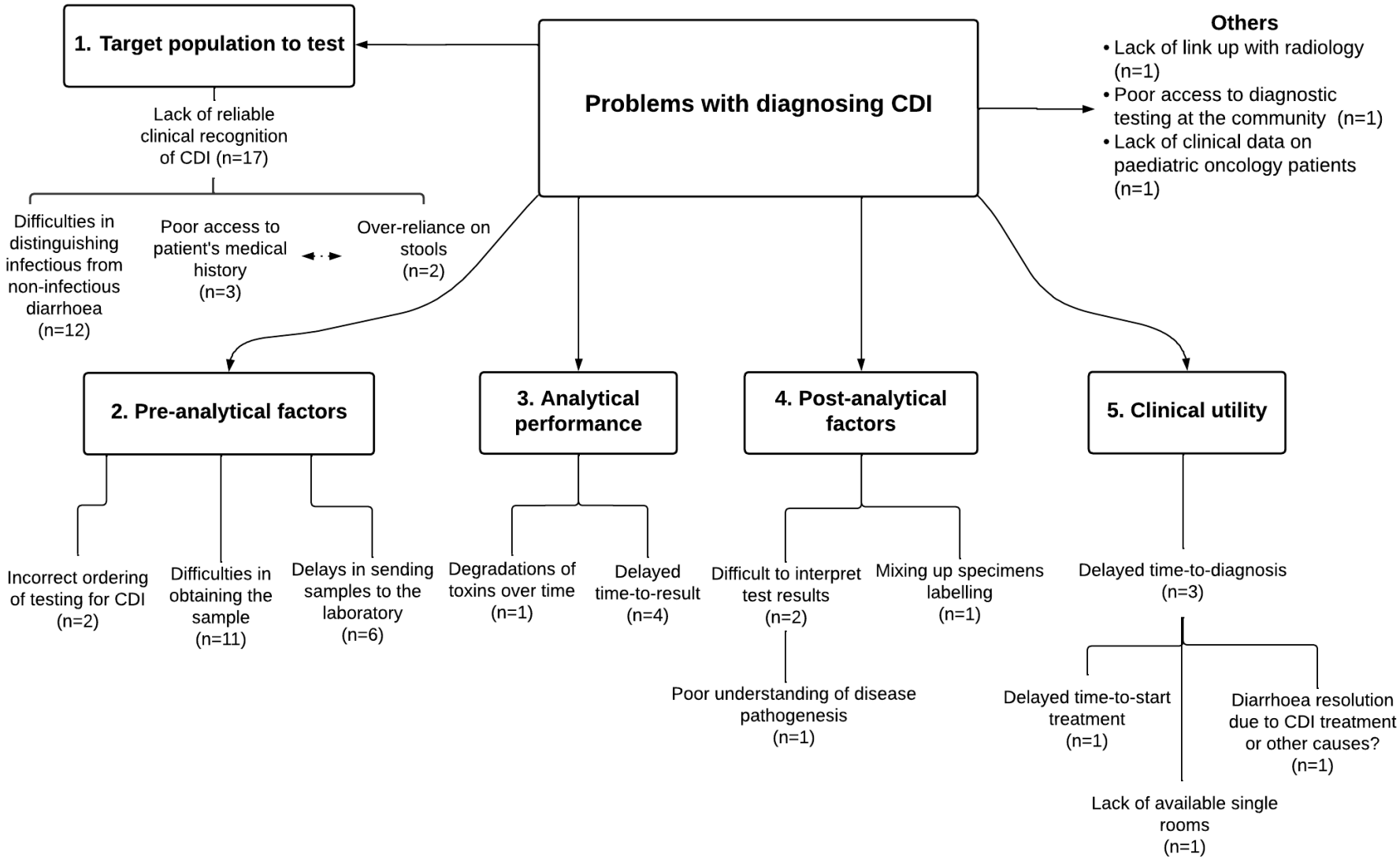
**Theme 4 – Post-analytical factors**

Two respondents highlighted challenges in understanding the results of multiple tests being run in a sequence, and as to how the combined results of a testing algorithm link to the disease pathogenesis. Specifically, one respondent noted a poor understanding of the disease pathogenesis (e.g. spores, carriage and active infection) among healthcare workers, whilst another mentioned that mixing-up of specimens might cause patients to receive an incorrect stool testing result.

**Theme 5 – Clinical utility**

Three respondents referred to delayed time-to-diagnosis as a common issue worsening patient health via delayed time-to-start treatment. One respondent stated that clinicians often rely on stool testing results before starting treatment despite patients having symptoms of CDI. Another respondent commented on the lack of an accurate and rapid POCT for CDI which, in turn, might lead to “a lot of side rooms needed at admission and no capacity” According to one respondent, delays in receiving test results for CDI might also cause confusion as to whether diarrhoea resolution was due to CDI treatment or to another undetected cause. Another respondent stated the lack of a reliable CDI diagnostic test for patients who remain symptomatic after treatment.

**Figure 4-10 Simplified schematic of the perceived issues with diagnosing patients with CDI**



#### 4.4.4.2 Problems with diagnostic tests for CDI

Every survey participant (n=48) was asked if there were any issues with current diagnostic tests for CDI available in their clinical setting, and what problems these limitations might cause.

Over a third of respondents reported problems with current diagnostic tests for CDI (n=19, 40%), as opposed to others reporting no issues with CDI diagnostics (n=12, 25%) or who were not aware of any limitations (n=17, 35%). Specifically, half of the ward-based clinicians did not know of any limitations to current diagnostic tests for CDI (n=16).

Among the respondents who reported issues (n=19), the most frequent limitations reported were low diagnostic sensitivity (n=9, 47%) and specificity (n= 9, 47%), followed by difficulties in interpreting test results (n=5, 26%) and long turnaround time (n=5, 26%) – see Figure 4-11. The most frequently reported consequences of the issues affecting current CDI diagnostics were “Delays in administering treatment to patients” (n=9, 47%), as well as “Inappropriate treatment for CDI” (n=8, 42%) and “Differences in case reporting across healthcare facilities” (n=8, 42%) – see Figure 4-12. Another common consequence of the limitations of current diagnostics for CDI was “Potential spread of infection” (n=7, 37%).

**Figure 4-11 Common perceived limitations associated with current diagnostic tests for CDI in absolute number (n) and proportions of respondents which selected that option within each subgroups (%).**

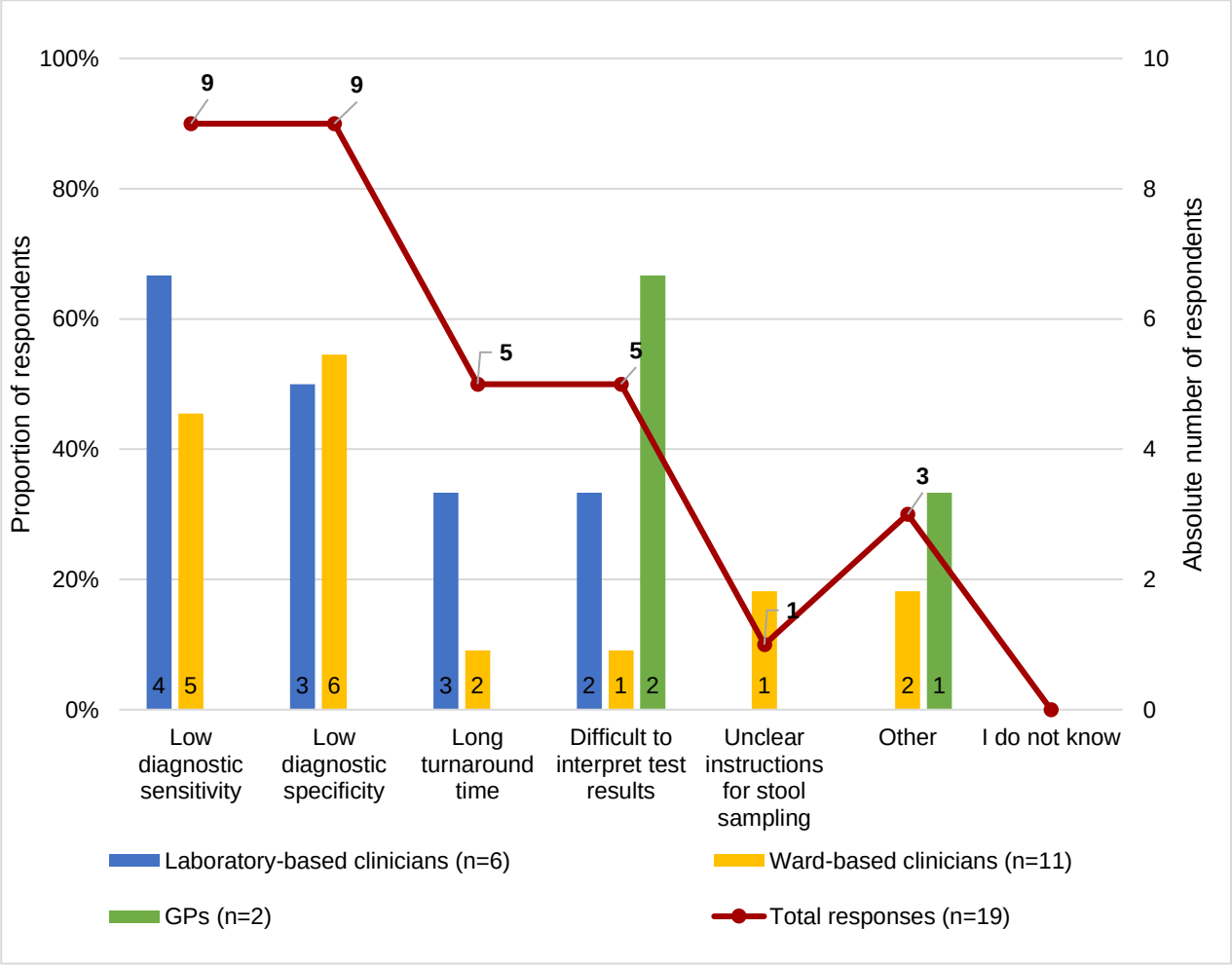
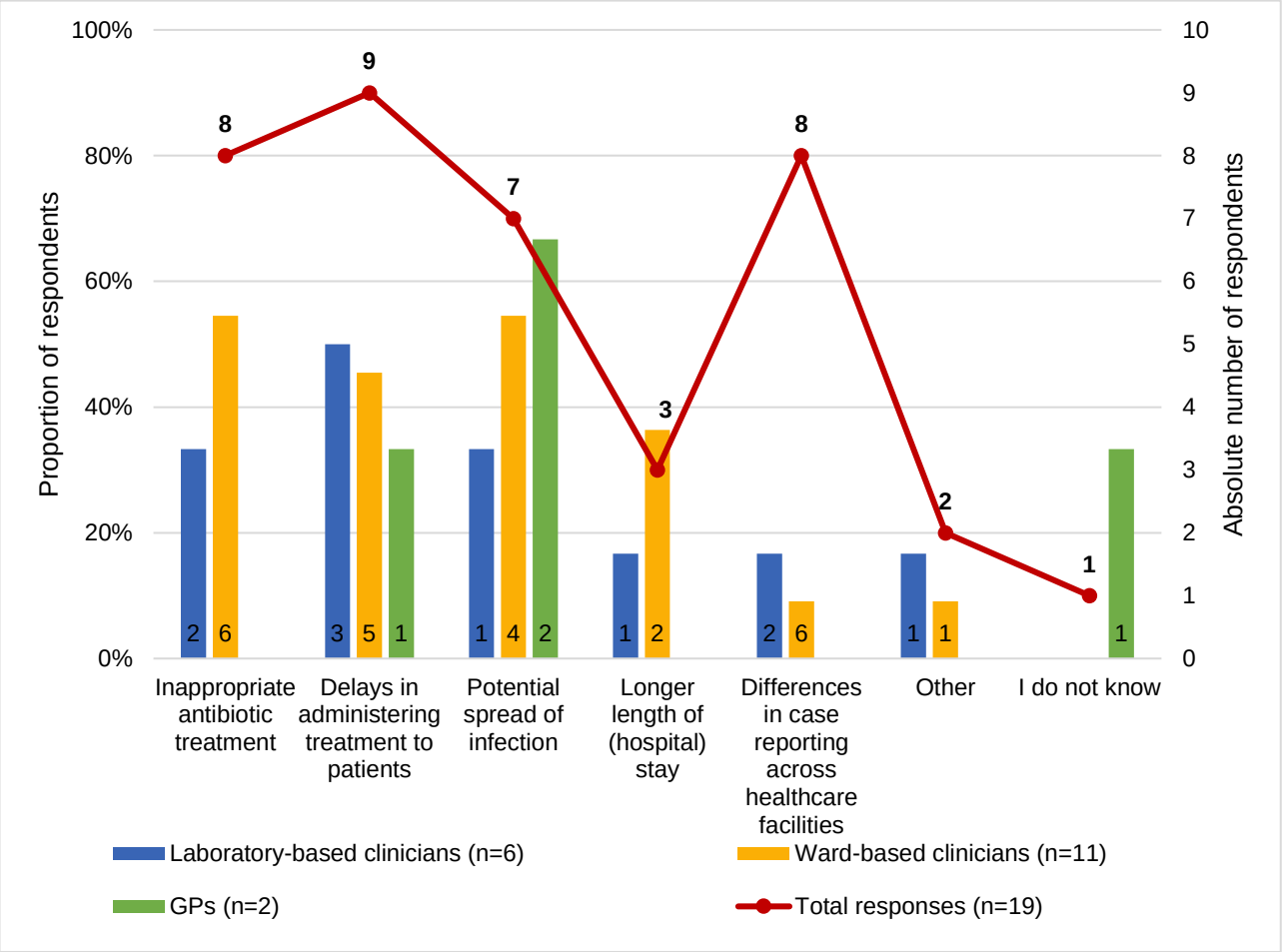


**Figure 4-12 Common consequences related to the problems associated with current diagnostic tests for CDI in absolute number (n) and proportions of respondents which selected that option within each subgroups (%).**




### 4.4.5 Room for improvement in current diagnostic tests?

Each of the survey participants (n=48) was asked to state if and how a better diagnostic test would solve any of the abovementioned issues with the diagnosis of CDI. Over a third of the respondents stated that new diagnostic tests for CDI might help to address issues affecting the diagnosis of CDI (n=19, 40%). Others reported insufficient room for improvement in current diagnostic tests for CDI (n=9, 19%), whereas the majority of respondents answered 'I do not know '(n=20, 42%).

Among those participants who stated that there was insufficient scope for improving new diagnostics for CDI (n=9), four respondents regarded the current testing system combining EIA toxin A/B and PCR as an efficient and cost-effective method of diagnosing patients with CDI, as a result of batch processing and a rapid turnaround time. Two respondents claimed that clinical symptoms of

CDI would mostly inform decisions as to whether to start antibiotic treatment for CDI and infection-control measures, regardless of laboratory diagnosis. One respondent mentioned that they were not aware of any limitations with existing CDI tests, whereas the remaining respondents did not answer (n=2).

Of the 19 respondents who noted issues with current CDI diagnostics, the majority agreed that new diagnostic tests for CDI might help to address current problems hindering the diagnosis of CDI (n=16). Respondents claiming that there was enough potential for improvement for new diagnostic tests for CDI (n=19) were asked how a new test could solve issues hindering the diagnosis of CDI. Most respondents answered this open question (n=18, 95%), with six themes identified – see Figure 4-13. The following sections describe each of the themes identified.

**Theme 1 – Target test population**

One respondent proposed a testing strategy which takes into account abdominal pain, which could help to reduce overreliance on volume and frequency of stools. Absence of stools should be considered as a sign of more severe infections – relying solely on stools' frequency for suspecting CDI might therefore miss some severe cases.

"Too much reliance on volume and frequency on stools – when absence of stool may be more indicative of serious infection" – *Clinical Scientist working in an English hospital*

**Theme 2 – Analytical performance**

Six respondents suggested that new tests should detect certain analytes, be it either toxigenic strains, or active production of toxins. One respondent proposed a new test detecting GDH, toxins A/B and toxigenic strains simultaneously.

"Higher sensitivity for *C. difficile* infection and for active toxin production would lead to greater confidence in the results" – *Infection Disease Doctor working in an English hospital*

*"*One that shows GDH, Toxin and toxin gene expression or non expression all in one!" – *Biomedical Scientist working in an English hospital*

One respondent also proposed that new testing options for *C. difficile* infection should also detect the bowel disease caused by CDI, or alternatively makers of inflammation or infections that are specific to *C. difficile* pathogen.

"Possibly a test that indicates bowel disease caused by C. diff [sic] i.e. a biomarker of damage to the bowel that is specific to c.diff [sic]." – *Consultant in Infectious Diseases working in an English hospital*

"[..] and include host specific markers of inflammation/ infection that relate to CDI and not other inflammatory bowel problems" – *Consultant in Medical Microbiology and Virology working in an English hospital*

One respondent suggested a test with the ability to rule out other possible causes of infectious diarrhoea as well as the presence of *C. difficile* pathogen.

"A quick and accurate test that can exclude cdiff, possibly a test that also looks for other causes at the same time" – *Consultant in Infectious Diseases working in an English hospital*

Four respondents also mentioned rapid time-to-diagnosis as a key feature new diagnostic tests should possess as a means of solving the issues affecting the diagnosis of CDI. In addition, one respondent suggested that having a standalone test could address some of the limitations hindering the diagnosis of patients with CDI. Alternatively, one respondent proposed a test with the possibility of random access (i.e. being able to test a single sample, rather than needing to test in batches).

**Theme 3 – Clinical validity**

Six respondents argued that new diagnostic tests for CDI with increased diagnostic accuracy would be desirable. Three respondents also suggested that new diagnostics for CDI should be able to differentiate between carriage of the *C. difficile* organism and the active infection.

"An assay that better differentiates carriage vs infection [..]"– *Clinical Scientist working in an English hospital*

**Theme 4 – Post-analytical factors**

Three respondents claimed that an increased ease of test result interpretation and ease of use would help to improve the diagnosis of CDI, or to reduce confusion around test results (n=1).

"Clarity in a single test would be ideal" – *Clinical Scientist working in an English hospital*

**Theme 5 – Infrastructural requirements**

One respondent pointed out that a higher availability of testing machines within laboratories might help to address some of the issues hindering the diagnosis of CDI, as well as access to more resources to support increased testing for CDI.

"Not the test itself but rather the availability of resources to test in the lab [sic]. Often labs [sic] do once a day runs" – *Infection Control Nurse working in an English hospital*

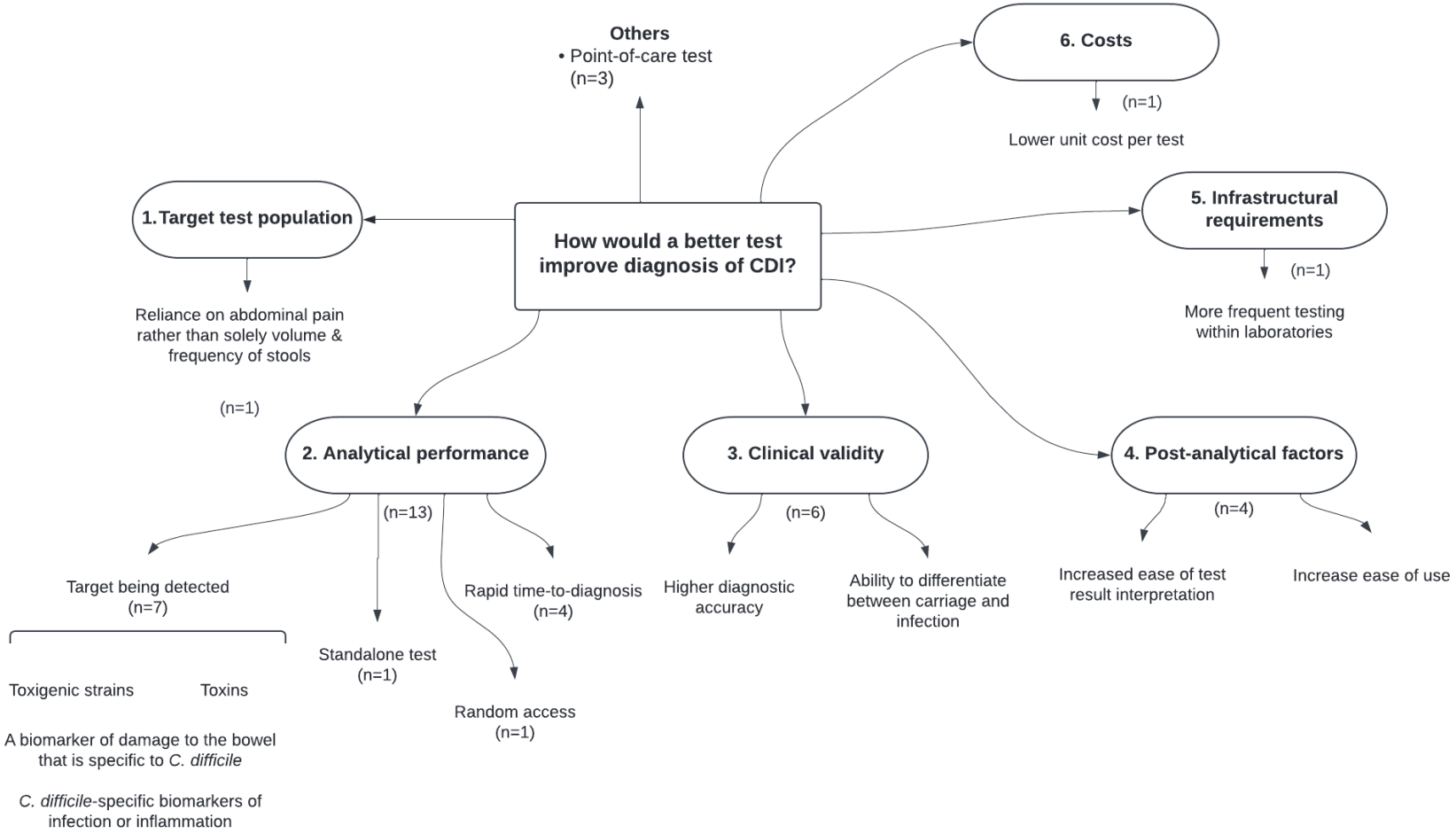
**Theme 6 – Costs**

One respondent also noted that new diagnostics for CDI should be less expensive.

**Others**

Three respondents suggested that a POCT could address some of the limitations hindering the diagnosis of patients with CDI.

**Figure 4-13 Desirable features new tests for *C. difficile* infection should possess to improve the diagnosis of patients suspected with CDI, sorted by main themes**



### 4.4.6 Ideal diagnostic test

Respondents who stated that there was sufficient room for improvement for new diagnostics for CDI (n=19) were asked to list ideal characteristics for new diagnostic tests, including: (i) type of sample; (ii) technique for obtaining the sample; (iii) method for transporting the sample; (iv) turnaround time of testing; (v) positioning of new test into existing care pathway; and (vi) test cost. The following ideal test characteristics were reported:

- **type of sample** – the most commonly reported ideal type of test sample for CDI was stool (n=15), followed by blood (n=4). One respondent alternatively suggested a rectal swab detecting skin microbes, instead of relying on stool samples. One respondent proposed “a sample more accessible than stool”;
- **technique for obtaining the sample** – rectal swabs were often recommended as an ideal technique for obtaining the samples (n=7), followed by stool collection (n=5). One respondent suggested that the ideal technique for obtaining samples should be “as non invasive as possible”;
- **technique for transporting the sample** – different suggestions emerged about this test feature. Three respondents suggested transporting swabs in tubes, a porter or using a chute system. Three respondents mentioned that the test should be a POCT without the need to transport the sample from the hospital ward to the laboratory;
- **turnaround time** – eight respondents proposed that the ideal turnaround time for new CDI diagnostics should be less than 1 hour, or within 2 and 6 hours (n=4). One respondent noted that rapid diagnosis of CDI (<2 hours) during emergency admission would be desirable, whereas another stated a preference for a point-of-care rapid test without quantifying the ideal turnaround time;
- **position in the care pathway** – several respondents recommended that an ideal test should either be point-of-care or ward-based (n=8), or run as soon as patients are identified as having diarrhoea (n=4);
- **acceptable cost** – several respondents reported that an acceptable unit cost of a new test would be less than £10 (n=7), followed by a price lower

than £30 (n=3) or £50 (n=1).The remaining respondents left this section blank or identified that they did not know how to answer (n=4).

Respondents also had the opportunity to list additional ideal features of new tests for CDI. Four themes emerged from this open question (n=11) – based on the key evidence domains underpinning the test evaluation methodology. See Figure 4-14 for more information on each theme and subtheme.

**Theme 1 – Target test population**

Two respondents expressed their interest for a new CDI diagnostic test specific to paediatric patients.

**Theme 2 – Analytical performance**

One respondent claimed that a test with the ability to distinguish between carriage of *C. difficile* pathogen and active toxin production might be preferable, whilst another suggested that an ideal new test for CDI should be able to detect different pathogens that might causing infectious diarrhoea. One respondent proposed a test detecting inflammatory markers in stools.

**Theme 3 – Clinical validity**

Four respondents would prefer new tests for CDI with very high diagnostic sensitivity and high diagnostic specificity.

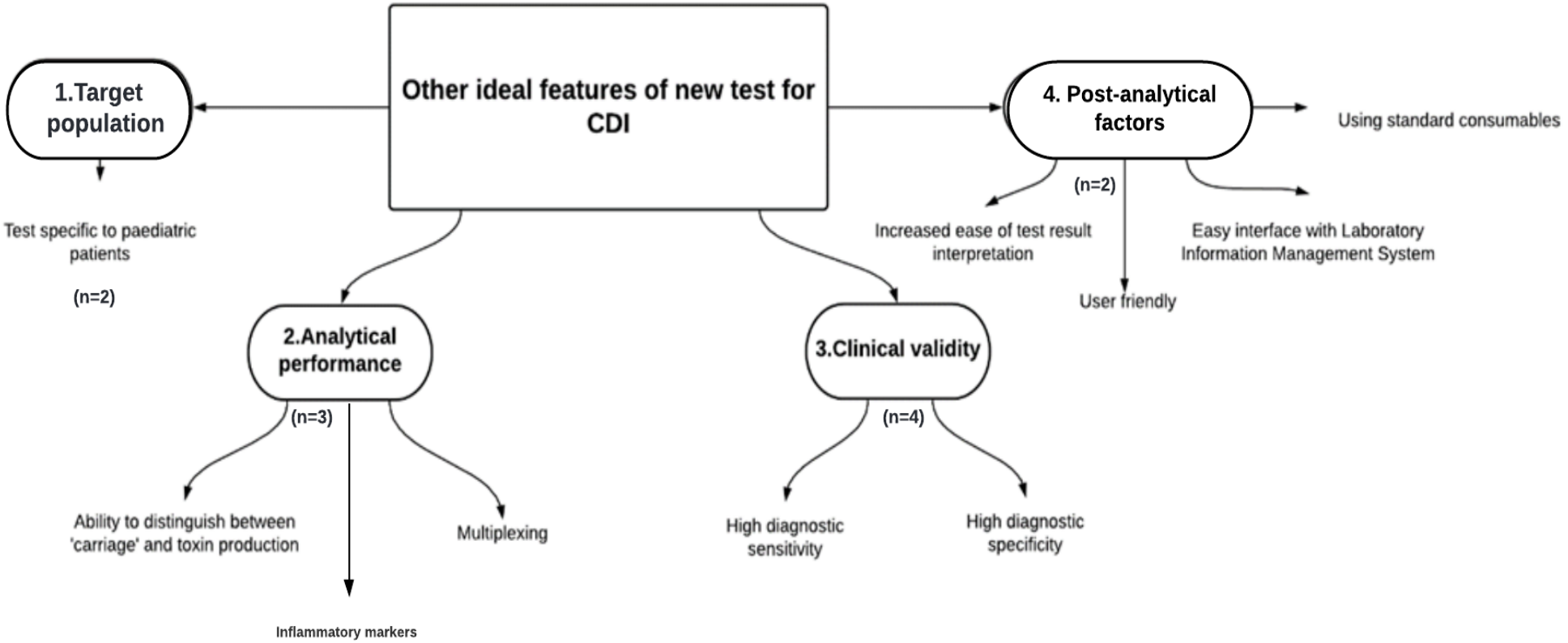
"would need very high sensitivity, specificity could be less if backed up by lab confirmation" – *Consultant in Infectious Diseases working in an English hospital*

**Theme 4 – Post-analytical factors**

One respondent emphasised the need for a test which is simple to run and does not require excessive training. Another respondent suggested that an ideal new test for CDI should be user friendly, with an easy interface with the Laboratory Information Management System (LIMS), enhanced ease of test results interpretation. For instance, results from a testing algorithm should be combined rather than receiving results from each testing option separately. Preferably, an ideal new test for CDI should use standard consumables.

"any algorithms Incorporated so result is clear with no interpretation necessary by the user. Interfaces with LIMS easily, uses standard consumables, user friendly" – *Clinical Scientist working in an English hospital*

**Figure 4-14 Additional ideal features new tests for CDI should possess, sorted by main themes and subthemes**



## 4.5 Discussion

To date, this is the first survey aiming to scope the unmet clinical need for new tests to diagnose CDI. A Europe-wide study using a Delphi method was conducted in 2014 to establish clinical priorities in the management and diagnosis of CDI patients (249), however that study did not explicitly define value propositions to drive innovation in new diagnostics for CDI. The results of this survey, on the contrary, describe the problems associated with the diagnostic tests for CDI, and the desirable features that a new diagnostic test for CDI should possess. These survey findings are therefore expected to be particularly useful for test manufacturers interested in developing new diagnostic tests for CDI that are ultimately 'fit for purpose'. In addition to this, findings from this survey describe the current diagnostic pathway for patients suspected with CDI and challenges within that pathway. These findings are therefore of interest to policy makers and relevant bodies in charge of developing national clinical guidelines to understand potential variations in the clinical pathway for CDI across the UK.

### 4.5.1 Mapping clinical care pathway for patients suspected with CDI

Watery stools and unexplained diarrhoea were the main symptoms found to prompt suspicion of CDI across ward-based clinicians – although the frequency and volume of stools were perceived as an inconsistent threshold for suspecting CDI. This survey confirmed that older patients were more frequently suspected of having CDI, and hence tested, whereas younger patients appeared to be underdiagnosed. This finding is consistent with several European epidemiological studies (250-253), such as the Longitudinal European Clostridium difficile Infection Diagnosis Surveillance Study (LuCID), EUCLID and COMBACTE-CDI studies.

A range of different CDI tests was used across different hospitals. EIA GDH and EIA for toxin A/B were the most commonly available diagnostic tests. This widespread adoption may be due to the lower cost of EIA, the possibility for batch testing, the quick turnaround time and ease of use (206, 214). Similar findings were found in a survey of 168 English acute NHS hospital trusts laboratories which indicated that EIA for toxin A/B was the most common diagnostic strategy for CDI (221).

Following the limitations of each standalone test for CDI, the majority of ward- and laboratory-based clinicians stated that they would run a testing algorithm – as recommended by European (222) and UK guidelines for the diagnosis of CDI (223). Based on the survey findings, a common testing algorithm was screening with EIA GDH, and then following-up with EIA toxin A/B ('Option 1') or with PCR ('Option 2'). The first test run as part of an algorithm rules out uninfected patients who can be managed as CDI negative cases (206). Testing options with high diagnostic sensitivity are therefore preferred as a screening test – as discussed in Chapter 3. The second step of a testing algorithm is used to rule out the presence of toxins in stools and thereby a second test with high specificity should be adopted – such as EIA for toxins A/B (222). Based on this logic, 'Option 1' and 'Option 2' appear to be suitable testing options for diagnosing patients suspected with CDI and ruling out non-infected patients.

Differences were found in the choice of tests included within testing algorithms, the order of the tests run (e.g. simultaneously or sequentially) and whether a confirmatory third test was requested. This finding is unsurprising given the lack of agreement around which reference standard to use for diagnosing CDI – be it either CCNA or TC – and the plethora of markers correlated with the *C. difficile* organism (212, 213), as described in Chapter 3. This extensive variability in testing protocols, however, might hinder the comparison of CDI incidence rates across healthcare facilities using different testing algorithms with varying diagnostic sensitivity and specificity (221).

While several factors were reported to inform the choice of testing algorithm – such as available testing equipment, the possibility of batch testing, laboratory's expertise and financial constraints (213), evidence on the comparative clinical utility and cost-effectiveness of different testing algorithms might better drive the choice of which testing protocol to run. Despite some exceptions (254, 255), there is currently a paucity of data on the short- and long-term impact that diverging testing algorithms for CDI might have on patient health. Further research is therefore needed to explore the comparative impact that different testing algorithms for CDI might have on patient health outcomes – (in terms of length of hospital stay, quality of life, and survival) – and the clinical care pathway – with reference to costs, in-ward transmission, and availability of single rooms.

Ward-based clinicians and GPs reported that they would consistently start infection-control measures (e.g. single room isolation, hand hygiene measures) upon suspicion and testing for CDI, and decisions around when to de-escalate these measures was mostly reported to be driven by resolution of symptoms. UK DH guidelines recommend de-isolating a confirmed CDI patient after 48 hours upon symptoms resolution (178), however some respondents would also consider other aspects such as the availability of single rooms and the full clinical picture of patients.

### 4.5.2 Issues affecting the diagnosis of CDI

The survey results indicated that most problems in the diagnosis of CDI revolved around the pre-analytical phase of testing – such as a lack of reliable clinical suspicion for CDI, and difficulties in stool sample collection. As diarrhoea is a side effect common to many conditions and treatments, there was uncertainty around who requires testing. Patients might be tested for CDI unnecessarily – because of a failure in identifying non-infectious causes of diarrhoea, for example; alternatively, their episodes of diarrhoea might go unnoticed due to a lack of clinical suspicion. Poor access to patient clinical history might prevent clinicians from understanding if, and which, antibiotics patients are currently taking. Although the accuracy of patients' self-reporting of antibiotic usage was found to be relatively reliable (256, 257), some survey respondents noted that patients might under-report their antibiotic prescriptions. Significant challenges were also reported with obtaining stools samples – especially in the context of incontinent or unwilling patients. This finding is not unique to CDI diagnosis – it is a widely reported fact that pre-analytical issues are the largest contributor to errors in clinical testing pathways (258, 259), as opposed to the issues arising in the analytical phase (i.e. sample analysis) where more stringent processes tend to be in place to standardise and automate as far as possible the testing procedures (259). It would appear therefore that for CDI diagnosis, as for most other diagnostic pathways, further work is required to improve the pre-analytical phase.

A commonly perceived limitation of testing options for CDI was poor diagnostic accuracy and long turnaround time. There were concerns that these issues led to delays to treatment or inappropriate antibiotic treatment. Despite the reported issues, when asked directly whether there were problems with currently available CDI diagnostic tests, most respondents replied by saying that they did not know

or that there were no problems. Of those respondents who stated that they did not know of any limitation with diagnostics for CDI, the majority were infection control nurses; it may therefore be conjectured that, as these healthcare professionals are not directly involved with the testing of patients suspected with CDI in the laboratory, they may therefore be less familiar with the potential limitations of current diagnostic tests when compared to laboratory-based clinicians. This same reason may have driven respondents to similarly report no specific problems with diagnostic tests for CDI – however the reasons underlying these subsets of answers are unknown based on the available data. Future work should focus on better investigating the reasons underlying these answers via interviews with the possibility to ask follow-up questions which, in turn, might yield richer and more comprehensive data.

Survey findings indicate three main areas R&D activities should focus on in relation to new diagnostic tests for CDI – namely enhanced patients' acceptability to sampling and sample handling, increased diagnostic accuracy and faster turnaround time. There is currently a paucity of data, however, as to whether early treatment for CDI, via early diagnosis, can provide a tangible clinical benefit to patients. A prospective time-series study, published in 2014, found that receiving a rapid diagnosis for CDI had a positive impact on patient clinical management (260). Further research, however, is needed to confirm the case for improved clinical outcomes resulting from early diagnosis of CDI.

### 4.5.3 Desirable specifications for new CDI diagnostics

Respondents suggested different desirable properties for new CDI diagnostics. A key requirement concerned the perceived need for new rapid POCTs for diagnosing CDI or with a quick turnaround time (e.g. less than 6 hours).

There was a lack of consensus on the desired specifications for a new diagnostic test for CDI across respondents – especially concerning pre-analytical and analytical factors. From an analysis perspective, stools were regarded as the ideal type of sample for new CDI diagnostics, however many respondents highlighted the difficulties in collecting stool samples from patients. Some respondents suggested a rectal swab as an alternative method of obtaining a sample. Others stated a technique that is as non-invasive as possible. If future tests could be developed to detect markers specific to the *C. difficile* organism based on more widely tolerated sample types (e.g. blood or saliva, as opposed

to stool), this would be of benefit. There is clearly a balance required here between what is plausible from a biological perspective and what is preferable to patients and healthcare professionals. Nevertheless, failure to consider patients' acceptability of sampling method when developing a test could lead to significant delays in sending samples to the laboratory – despite the diagnostic test itself being rapid. There was also a lack of agreement among respondents on whether the test should detect toxigenic strains or active toxin production in stools. This uncertainty reflects an ongoing conundrum hindering the understanding of CDI pathogenesis (226).

In addition to CDI-specific analytes being detected, several respondents suggested developing a test which measures bowel disease caused by CDI, or alternatively other GI pathogens that might be causing diarrhoea. Although multiplex GI panels detecting *C. difficile* toxigenic strains and other GI pathogens are already on the market, their use for diagnosing CDI is under debate as detection of toxigenic strain alone is not indicative of CDI (261) – as discussed in Chapter 3.

Several respondents also recommended high diagnostic accuracy as a key feature for new CDI diagnostics. While the survey did not purposefully ask respondents to quantify the desirable diagnostic accuracy for new CDI diagnostics, the *de novo* early economic model outlined in Chapter 5 and Chapter 6 aims to address this question.

### 4.5.4 Study strengths and limitations

This online survey had both strengths and limitations. The use of conditional branching questions allowed customisation of the questions being asked depending on the respondents' field of expertise, clinical setting of interest and knowledge regarding testing options for CDI. This helped to reduce the number of questions asked while also gathering more relevant data. The extensive validation process undertaken ensured the face and content validity of the survey.

This survey, however, was based on a non-probability convenience sample of UK ward- and laboratory-based clinicians and GPs who saw the survey advertised on social media or emails and volunteered to participate. The generalisability of the survey findings may therefore be limited due to a risk of *volunteer bias* – which is common among open surveys (238). As the research team was not in control

of the selection process, only respondents who were interested in filling the survey participated in the study (i.e. *self-selection*) (262). The current survey design therefore failed to prevent multiple entries from an individual respondent – although the likelihood of that occurring is expected to be low. Future versions of this survey could address this issue by excluding responses submitted within an atypical timestamp (238).

An additional limitation of the survey design lies in the lack of screening questions to certify that the survey participants fit a pre-specified set of criteria (i.e. having experience of diagnosing CDI and being familiar with diagnostic tests for CDI). As the survey was open to anyone with access to the survey link, there is a chance that some survey respondents were inaccurate information about their professional status and experience with CDI tests. This seems unlikely given the quality and detail of responses received, however there is a chance that this had an impact on the validity and representativeness of the survey results. Future iterations of this survey could include a set of screening questions at the end of the information sheet or when asking respondents about their details (e.g. job title, specialty, location) to rule out those whom did not meet the pre-specified criteria. For example, asking respondents upfront if they were familiar with the national guidelines for diagnosing CDI, or to correctly identify the diagnostic tests for CDI among a list of diagnostic tests for several conditions, or if they have had experience in diagnosing patients suspected with CDI could have helped to ensure that only those respondents meeting the criteria would fill in the survey.

In addition, questions being asked, or items within multiple-choice questions, were not randomised thereby potentially leading to *question order bias* (238). The software where the online survey was programmed (i.e. Online Surveys) did not allow for randomising items within multiple-choice questions, and it was decided against varying the order of questions as this was heavily dependent on adaptive questioning.

Findings from this survey are generalisable to ward- and laboratory-based clinicians working in English hospitals. This study also did not capture the views of other relevant stakeholder groups in the field of CDI – such as hospital managers, patient and industry representatives. Defining unmet clinical needs is a multidisciplinary exercise which requires the input from as many stakeholder groups as possible. For example, the systematic review (Chapter 2) found that

the patient perspective has so far only played a minor role in typical TPP development process. Ensuring that clinical unmet needs are derived from both clinician and patient's views and preferences is of paramount important to ensure that the end-user perspective is taken into consideration. This will, in turn, maximise the likelihood that a novel test is 'fit-for-purpose' and provides clinical benefits to patients. To address this limitation, future focus groups, or alternatively an additional survey, could help to gauge the views of patients and those stakeholders to whom the survey was not originally sent out. This, in turn, could help to reach a broader consensus on the clinical priorities for CDI diagnostics being identified in the survey.

A final limitation of this study is the limited sample size (n=48); in particular the low number of GPs responded to the survey (n=4). This may be due to the convenience nature of the sampling – testing for CDI in the community is not as common (263) and therefore GPs may not have been as drawn to the survey invitation. The original dissemination of the survey was planned just as the COVID-19 pandemic was gaining force (April to August 2020). This was an extremely busy period for all healthcare professionals, particularly those working in the field of infection. Attempts to mitigate the initial limited sample size were made by having additional rounds of dissemination, with partial success. To obtain a larger survey response however, future studies of this kind will likely need to wait until the pandemic effect has significantly subsided, or be able to invest substantial efforts into wide-scale recruitment.

While the sample size of this study is considered limited – especially with respect to surveys, it constitutes an improvement in the context of published clinical needs assessment studies based on focus groups which tend to involve a much lower number of healthcare professionals (20). The survey results were able to capture notable differences in the diagnostic pathways and testing options for CDI across UK hospitals, and highlight common challenges faced by a wide range of healthcare professionals with varying expertise in the diagnosis of CDI. Findings of this survey are therefore expected to be of great interest to test developers, as well as clinicians and policy makers in this setting, to help them understand the current testing practices and perceived challenges with diagnosing CDI.

## 4.6 Chapter summary

- This chapter presented the development and results of an online survey of UK healthcare professionals, which aimed to understand current clinical practices for diagnosing CDI and to assess if there is an unmet clinical need for new diagnostic tests for CDI.
- Forty-eight participants completed the survey, most of whom were infection control nurses (n = 21). A variety of testing algorithms were reported as being used across different hospital laboratories, with running EIA GDH and EIA toxin A/B being the most commonly reported testing algorithm.
- The primary issues identified with current diagnostics for CDI concerned difficulties in identifying who requires testing, and problems with sample collection (i.e. difficulties in obtaining an adequate stool sample). Test turnaround time was also highlighted as an issue – thereby a perceived need for new rapid POCTs for diagnosing CDI was found.
- Key identified requirements for a new CDI diagnostic included: (i) quicker turnaround time, (ii) less invasive sample requirements, and (iii) high diagnostic accuracy.

This online survey did not aim to elicit respondents' views on the minimum and desirable diagnostic accuracy for new diagnostic tests for CDI. To this end, Chapter 5 presents a *de novo* early economic model which aims to evaluate the cost-effectiveness a hypothetical rapid test for CDI. Chapter 6 outlines the analysis conducted to estimate the minimum diagnostic accuracy and maximum unit price for a hypothetical rapid test for CDI based on clinical and cost-effectiveness considerations.

# Chapter 5

# Early economic model comparing a new hypothetical rapid diagnostic test for CDI against standard care – structure and parameterisation

## 5.1 Chapter outline

In Chapter 3, the clinical background of the case study was provided, alongside an overview of the typical clinical pathways for patients suspected with CDI. This chapter describes the structure, development and parameterisation of a *de novo* early economic model comparing the cost-effectiveness of a hypothetical diagnostic test for CDI to current clinical testing practice at LTHT. Key findings from a rapid literature review of decision models evaluating the cost-effectiveness of diagnostic tests for CDI are outlined in section 5.2. An overview of the structure of the model is then provided (section 5.3), including: (i) rationale for the modelling approach (section 5.3.2); (ii) details of the conceptual structure of the model and main clinical assumptions (section 5.3); (iii) an analysis of the simulated disease spread (section 5.3.4); (iv) parameterisation of the model (section 5.3.5); (v) an outline of the model outputs (section 5.3.6); (vi) details of model implementation (section 5.3.7); and (vii) an overview of model validation process (section 5.3.8). Model analysis and the estimated performance specifications are subsequently presented in Chapter 6.

## 5.2 Background

To inform the development of the model, a rapid literature review and critical appraisal of decision models evaluating the cost-effectiveness of diagnostic tests for CDI was performed. Full details of the methods and results can be found in Appendix J and K, respectively. Eight studies, published between 2012 and 2020, were included (235, 264-270). Common modelling approaches were identified, alongside some limitations of existing models. The mechanism of impact of CDI diagnostics on patient health was typically modelled by triggering infection-control measures (e.g. single room isolation) and/or administration of antibiotic treatment regimens for CDI.

When critically appraising the existing models, some common methodological limitations were found. Firstly, all models failed to capture further diagnostic activity following the receipt of an initial CDI result. In clinical practice, clinicians would usually request subsequent testing upon receipt of a negative test result to rule out other causes of infectious diarrhoea, or alternatively, re-test for CDI in cases where symptoms persist (266). Every included study, however, assumed that these patients would either be discharged from the hospital or exit the model. Secondly, for the models where a CUA was conducted, the evidence base underpinning the utility weight parameters for CDI was weak. Often, health utilities for non-infectious diarrhoea were adopted as proxies for CDI (264, 265, 268) – such as ulcerative colitis, chronic inflammatory bowel disease, and prostate cancer. As a consequence, the impact of CDI on patients' health quality of life and QALYs may have been underestimated as non-infectious diarrhoea is not as severe as CDI. This might then lead to biased conclusions on the expected cost-utility of diagnostic tests for CDI. A third limitation, which relates to the ability of the standard economic evaluation framework to capture the key advantages and disadvantages of the intervention, was the oversight of capacity considerations when evaluating the cost-effectiveness of CDI diagnostics. None of the included studies captured the consequence of testing strategies for CDI on the availability of single rooms to isolate patients suspected or confirmed with CDI – except one stochastic system dynamic (SD) model (270). While several studies explored the impact of receiving an incorrect diagnosis on the implementation of infection-control measures (e.g. single room isolation) (235, 264, 267, 270), the *availability* of single rooms was seldom accounted for. Given that the availability of single rooms is a key structural aspect that impacts on hospital workflow and subsequent infection spread, this is a key limitation of past modelling studies.

The review findings were used to inform the structure and parameterisation of the early economic model comparing the cost-effectiveness of a hypothetical diagnostic test for CDI discussed in section 5.3.

## 5.3 *De novo* early economic model

The sections below give an overview of the *de novo* early economic model (henceforth referred as 'the model').

### 5.3.1 Scope of the model

The aim of this research is to develop an early economic model to help inform key TPP characteristics for a new rapid diagnostic test for CDI, based on the clinical utility and cost-effectiveness outputs of the EEE – including INMB. Specifically, this analysis is focused on developing a decision-analytic model to:

- explore the impact that a POCT could have on infection control infrastructure (e.g. availability of isolation rooms), clinical decision-making, infection spread (e.g. new secondary CDI cases in general ward) and costs; and
- identify the necessary properties of a POCT for CDI to be cost-effective compared to standard care, from a UK NHS perspective (focusing on minimum diagnostic sensitivity and specificity, turnaround time and maximum cost for the test).

### 5.3.2 Choice of modelling technique

Different modelling techniques are available to conduct economic evaluations. An overview of common modelling techniques (including decision tree models, Markov models, DES models, SD models, and agent-based simulation models) is provided in Appendix A.

Selecting an appropriate modelling approach is contingent on the decision being evaluated, as well as the key mechanisms that needed to be captured within the model (e.g. interactions between individuals, timing of events, resource constraints) (271, 272) (see Figure 5-1). Several authors have published guidance on this topic (271, 273-277). Key factors include (278): (a) whether individual characteristics of the modelled population need to be captured or not (i.e. individual-level vs. cohort-level modelling); (b) whether the precise timing of events needs to be captured; (c) whether interactions between individuals are relevant for the decision at hand; (d) whether resource constraints should be captured or not; and (e) what the modeller's resources are (e.g. project constraints, software costs and modelling expertise).

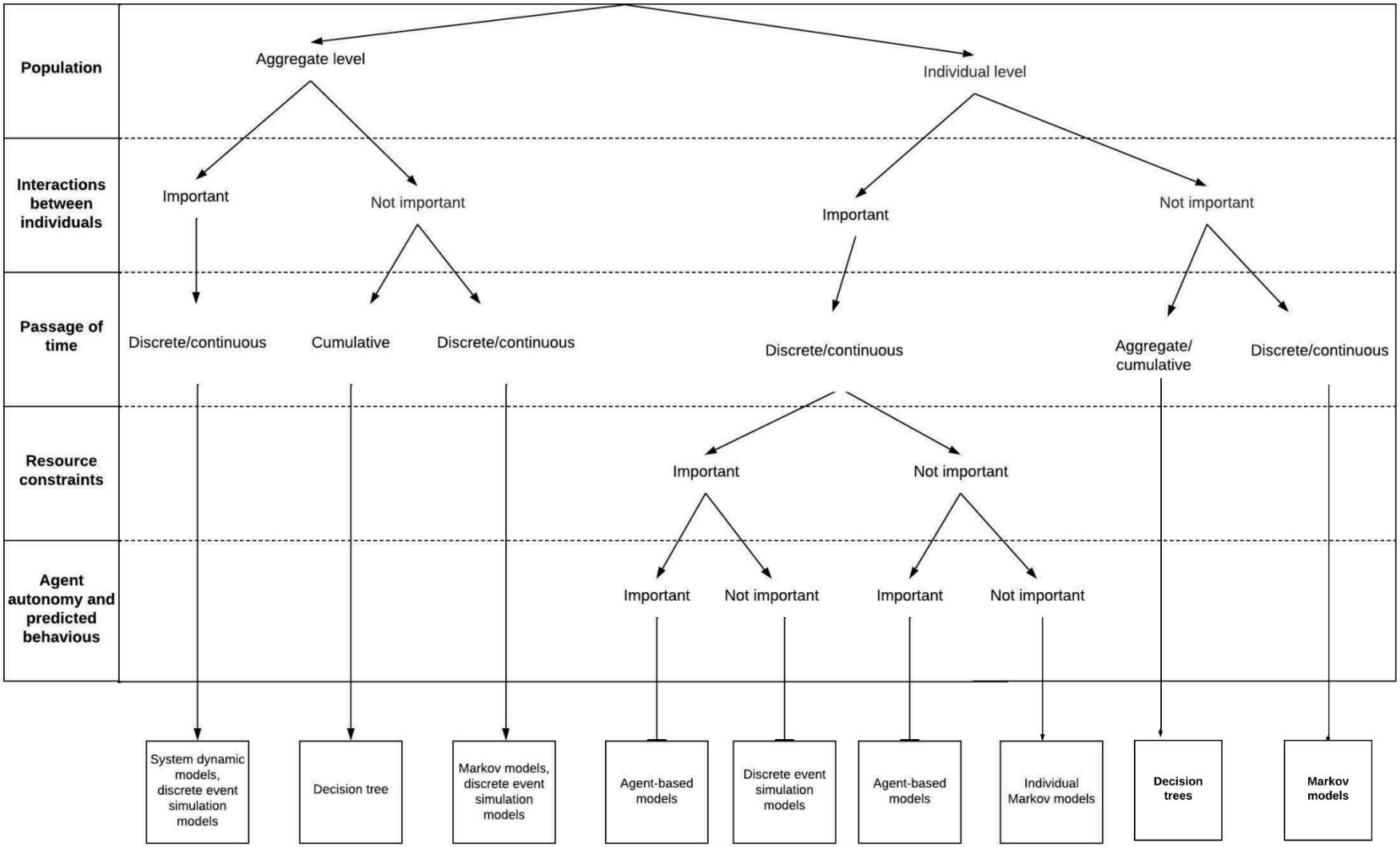



**Figure 5-1 A flowchart summarising the key requirements for a decision model. Figure adapted from Stahl, J.E. (275).**

Ideally, the model comparing the cost-effectiveness of a hypothetical rapid diagnostic test for CDI should capture:

- **individual patient characteristics** – clinical decisions for patients suspected and confirmed with CDI (e.g. antimicrobial treatment decisions, when to de-escalate infection control measures) take into consideration many individual factors such as number or recurrences, disease severity and patient risk factors. To explore how a new POCT could impact clinical decision-making for CDI patients, the model should account for individual patient factors rather than assuming homogeneity within the patient cohort as per cohort-based models (e.g. decision trees, Markov models) or aggregate models – such as SD models[15];
- **patient history** – according to clinical practice, patients receiving a negative CDI test result might be tested again for CDI or other GI pathogens if symptoms persist (181). In addition, the duration of antimicrobial treatment for CDI is dependent on resolution of symptoms and therefore a clinician might check if a given patient is symptomatic repeatedly over the patient's hospital stay. The ability to retain patient history information or past events is therefore desirable as it allows to determine the future patient trajectory within the clinical pathway, disease progression and associated resource requirements based on patient's current status (e.g. test result, resolution of symptoms). This enables evaluation of capacity implications as it is possible to account for what resources each individual patient needs (e.g. single rooms);
- **time-to-event** – the timing of events in the context of CDI diagnosis (e.g. when test results are received, when to start an antibiotic treatment regimen, when and for how long a patient remains in presumptive isolation) is expected to have an impact on the availability of single rooms, costs and patient health. The ability of capture the passage of time – either continuously as per SD and agent-based simulation (ABS) models, or at discrete timepoints as per DES models, is therefore desirable to more accurately estimate patient health outcomes, costs and resource constraints;

[15] An alternative name for SD models is ordinary differential equation (ODE) models.

- **interactions between individuals** – interactions between patients, clinicians and visitors are the main vector for disease transmission in this case study, as patients positive to CDI or colonised can spread the pathogens to other individuals through shedding common surfaces (e.g. bathrooms, wards) (279). Interactions between individuals also impact the infection-control infrastructure as patients in isolation who are considered lower priority from an infection-control perspective might be de-isolated to release single rooms for patients with higher priority. DES, ABS and SD models allow the user to capture interactions between individuals over time, as opposed to Markov models and decision trees which do not have this feature; and
- **resource constraints** – in the UK, inpatients suspected with CDI are typically isolated in single rooms while awaiting test results to prevent in-hospital transmission, until confirmation of non-infectious diarrhoea (181). Core to this decision model is therefore exploring what impact a hypothetical rapid test could have on the infection control infrastructure and clinical decision-making (e.g. start isolation, obtaining and transporting samples to the laboratory).

Based on the key requirements for the model presented above, decision trees and Markov models can be ruled out as a suitable modelling technique for this case study as: (i) they do not allow for interactions between individuals; (ii) they do not capture resource and/or capacity constraints; (iii) they do not accurately simulate the passage of time; and (iv) they do not retain patient past information and history.

Three modelling techniques are usually employed to simulate healthcare-associated infections – namely DES, ABS and SD modelling (279). Based on an assessment of the different capabilities of DES, ABS, and SD models, DES was selected as the technique of choice for the model. The primary advantages of DES modelling in this case are: (i) the ability to capture individual patient characteristics (e.g. age, duration of symptoms and risk of experiencing a certain event) and history in the model (possible with both DES and ABS modelling, as opposed to SD modelling); (ii) the ability to accurately measure the timings of events, and their associated costs and outcomes – as within DES models transitions are evaluated as they occur rather than at fixed time cycles as within

Markov models); (iii) the ability to capture a sequence of hospital processes and activities patients suspected with CDI undergo within the care pathway (possible with both DES and ABS modelling technique only) – rather than simulating the predicted behaviour of patients and healthcare workers as autonomous agents as per ABS models[16]; and (iv) the ability to capture capacity constraints for scarce resources (e.g. single rooms) and queues in the system (common between constrained-resource DES and ABS modelling, as opposed to SD models).

Whilst DES is not the optimal approach for modelling infection spread compared to SD or ABS models (279), it is nevertheless able to indirectly estimate secondary infections, which is the most common approach taken in past models to capture CDI spread (235, 264). DES can provide useful information on secondary infections in an efficient and static way, without requiring the same level of data on infection transmission needed for a SD or ABS model - which in the context of a hypothetical test would not be expected to be available. Details on estimation of the secondary cases of CDI can be found in section 5.3.4, while the potential impact of this limitation is discussed in Chapter 6.

It is also of interest to highlight that a similar modelling approach, utilising a DES model, was adopted within the recent early economic model of COVID-19 POCTs published as part of a NICE-commissioned project (164). While the authors did not explain the reasons as to why this modelling technique was chosen, building a DES model allowed them to: (i) capture the infection spread of COVID-19 between patients and healthcare professionals; (ii) simulate recurring events (e.g. patients can be retested in case of an unclear test result); and (iii) assess the impact of testing turnaround time on the ward bays where patients suspected with COVID-19 were assumed to await test results.

[16] As this model aims to evaluate the impact of a hypothetical test on the clinical pathway, incorporating agent autonomy and predicted behaviour was not expected to provide meaningful information for the clinical stakeholders in the context of TPP development. ABS models can therefore be ruled out as an appropriate modelling technique for the decision at hand.

### 5.3.3 Model structure

A deterministic, stochastic[17], resource-constrained DES model was developed in SIMUL8 (SIMUL8 Corp, Boston, MA) to map the flow of patients presenting with CDI symptoms through the hospital setting, accounting for key capacity constraints (i.e. availability of single rooms, laboratory schedule for processing samples). The model is static, in the sense that it does not capture directly the infection spread of CDI within the hospital. Instead, the model indirectly estimates the number of new secondary infections of CDI due to a lack of infection-control measures when confirmed CDI cases remain in general ward due to a limited availability of single rooms (see section 5.3.4 for more details).

As clinical practices for diagnosing CDI in the hospital currently vary across the UK (based on the findings from the survey presented in Chapter 4), the pragmatic decision was made to base the model structure on a single centre, but whilst ensuring sufficient model flexibility to explore alternative hospital configurations via sensitivity and scenario analyses (e.g. assuming a different volume of patients arriving and number of single rooms). The model was therefore built to reflect current clinical pathways in place at LTHT (281, 282) – a medium to large sized hospital which sees roughly a million patients annually.

Consultations with clinical experts based at LTHT and a review of local clinical guidelines, in addition to the studies identified in the rapid literature review outlined in Appendix K, informed the model development (181, 281-284). Specifically, information on key processes underlying the LTHT laboratory workflow for diagnosing CDI and clinical management for patients suspected with CDI, outlined below, were derived from consultations with an LTHT Principal Clinical Scientist and a Clinical Research Fellow, respectively.

Figure 5-2 represents a simplified schematic of the model structure. The upper branch of the model depicts the movement of patients through the hospital, while the lower branch depicts the testing processes undertaken on each patient's test sample. Further details on the clinical assumptions underpinning the model, and

[17] Stochastic models are subject to *first-order* uncertainty which reflects "the fact that individuals facing the same probabilities and outcomes will experience the effects of a disease or intervention differently [..] (e.g. the first patient in a sample might respond to a treatment but the next one may not)" (280). Section 5.3.7.2 outlines more details on how the uncertainty was handled in the model.

each of the subcomponents of the model are provided in the following sections. Full model coding and technical documentation is available upon request.

The model was developed in accordance to the Strengthening The Reporting of Empirical Simulation Studies (STRESS) guidelines (285) and Consolidated Health Economic Evaluation Reporting Standards (CHEERS) (286) (see Appendix M for detail).

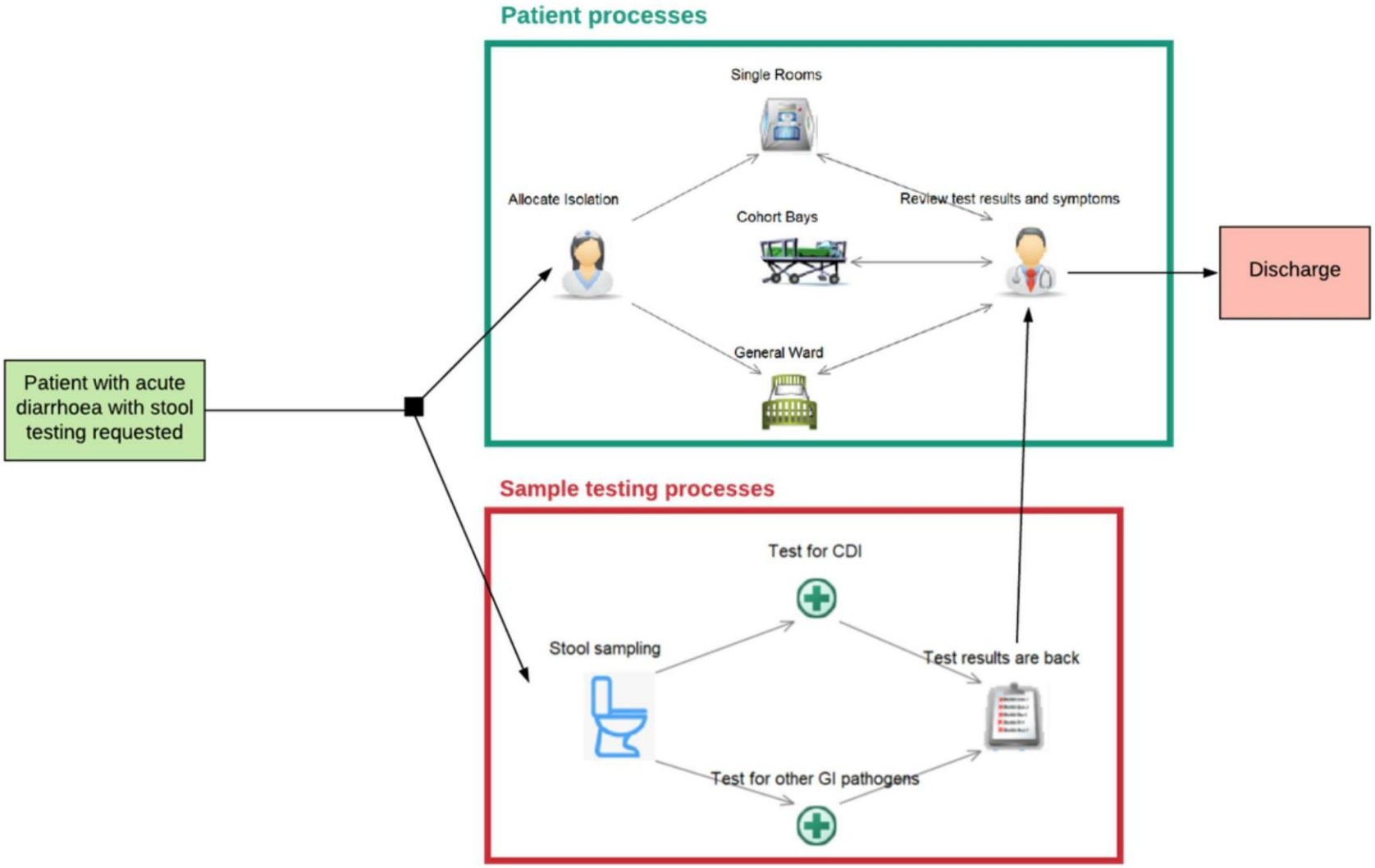



**Figure 5-2 Simplified schematic of the early economic model for a new hypothetical rapid diagnostic test for CDI**

#### 5.3.3.1 Clinical assumptions

Key clinical assumptions (CA) underpinning the model are summarised in Table 5-1. These assumptions were informed by a series of consultations with clinical experts (including a Principal Clinical Scientist and a Clinical Research Fellow at LTHT).

**Table 5-1 Key clinical assumptions of the model**

| Clinical Assumption |
|---|
| **CA 1.** The monthly demand for stool testing was assumed to be constant and independent from seasonal change or other external factors. |
| **CA 2.** There is no risk of co-infection between GI pathogens and CDI. |
| **CA 3.** Isolation in single rooms is required for patients positive to other GI pathogens. |
| **CA 4.** When single rooms and cohort bays are at full capacity, patients confirmed with infectious diarrhoea remain in general ward until hospital discharge. |
| **CA 5**. A rapid multiplex GI panel with assumed perfected diagnostic accuracy is run to detect multiple pathogens – separate to tests for CDI.<br>Perfected diagnostic accuracy was assumed for the multiplex GI panel as the focus of this decision model lies in evaluating testing strategies for CDI, rather than other GI pathogens. |
| **CA 6.** Clinicians consider the continuation or resolution of diarrhoea as the only symptom for CDI – without assessing whether the frequency of stools is improving or worsening. See Chapter 4 for more information. |
| **CA 7.** Clinicians are assumed to pre-emptively administer treatment for CDI (i.e. 'empirical treatment') prior to receipt of test results if a patient presents with severe symptoms of CDI – as per current clinical practice at LTHT. |
| **CA 8.** Clinicians fully adhere to test results when deciding when to start or stop administering antibiotic treatment for CDI. |
| **CA 9.** Early treatment, via early diagnosis, is assumed to have no impact on patient survival, risk of disease recurrence and long-term quality of life due to paucity of data. |
| **CA 10.** Patients with a true-positive (TP) test result for CDI are assumed to recover to full health without risk of disease recurrence at the end of the antibiotic treatment regimen. |
| **CA 11.** Patients with a FN test result for CDI do not experience further disease recurrences. |
| **CA 12.** Patients with a FP test result for CDI are assumed to remain symptomatic at day 10 of treatment and receive treatment for 4 additional days – as per current clinical practice at LTHT.<br>For more information see section 5.3.3.6. |

#### 5.3.3.2 Starting point

In-hospital adult patients with an initial episode of acute diarrhoea for whom clinicians have requested stool testing for CDI enter the model.

Patients are moved into presumptive isolation, while their stool samples are tested simultaneously for CDI and other GI pathogens which could cause diarrhoea. Stool testing and presumptive isolation happen simultaneously. The simulation software has the functionality to divide a single *entity* (i.e. patient) into two parts (called *batching* in SIMUL8) which share the same individual-level information (e.g. disease prevalence, time to enter the model), and to re-combine those two parts (called *components*) of an entity at a later event (i.e. once test results are back). This enables the user to simulate different events happening simultaneously to a single entity. This approach is used in the model to track: (i) which single isolation room a patient enters, depending on the current availability of single rooms, and (ii) the various processes of the testing pathway that each individual's test sample undergoes (e.g. sample preparation, setting the machine, reviewing test results). The division of entities into their two respective components is undertaken at the start of the model. Once the patient's test result is received, the two components are recombined and information on an individual patient's health is updated.

#### 5.3.3.3 Escalation of infection-control measures

LTHT clinical guidelines recommend placing a patient into presumptive isolation within two hours of suspicion of infective diarrhoea to reduce the risk of nosocomial transmission within the general ward (281, 283). Two locations are available for patients suspected with infectious diarrhoea depending on current capacity constraints:

- **single rooms**: only one patient can enter a single room with no potential of infecting others. If there is a confirmation of CDI, the patient remains in isolation until the end of their hospital LOS; and
- **general ward**: when no single rooms are available, patients remain in the general ward while waiting for test result with a higher potential for infection transmission. The general ward is set to an unlimited capacity in the model, to host as many patients as possible, where necessary.

Cohort bays are available for patients confirmed with infectious diarrhoea only – in the absence of available single rooms. At LTHT, patients with the same detected infection can be grouped together in four cohort bays (each with a maximum capacity of 6 patients).

#### 5.3.3.4 Testing pathway

Upon sample collection, the sample is shipped to the laboratory or tested within the ward depending on the testing strategy under evaluation. The following features are common to both testing strategies being evaluated: (i) test turnaround time represents how long it takes to yield test results after having obtained the sample to test; (ii) the sample is simultaneously tested for CDI and other GI pathogens to rule out other causes of infectious diarrhoea using a rapid multiplex GI panel with perfected diagnostic accuracy; and (iii) upon receipt of final diagnosis, the test sample is matched to the corresponding patient within the patient isolation pathway, to update information on patient's health following test results.

##### 5.3.3.4.1 Standard care testing for CDI

Details of the LTHT laboratory clinical pathway for CDI, outlined below, were derived from consultations with a Principal Clinical Scientist based at LTHT. At LTHT an on-site laboratory routinely processes stool samples at 10am and 4pm. Samples arriving before these time points will wait before being processed.

A two-step testing algorithm is currently run at LTHT for patients suspected with CDI (Figure 5-3). Each sample is initially screened with TECHLAB ® C.DIFF CHECK$^{TM}$-60 GDH EIA to detect the presence of *C. difficile* organism. If a sample is negative to GDH EIA, it is possible to exclude CDI.

If a sample is positive on GDH screening, a Cepheid ® Xpert *C. difficile* PCR and CCNA are used simultaneously to detect toxin genes related to CDI or free toxins in stools, respectively. PCR testing yields results quickly (e.g. 43 minutes), whilst CCNA ultimately confirms the presence of free toxins within two days. Final diagnosis is confirmed upon receipt of CCNA results. Table 5-2 gives an overview of the possible diagnosis, interpretation of test results and clinical

recommendation based on both PCR and CCNA test results.

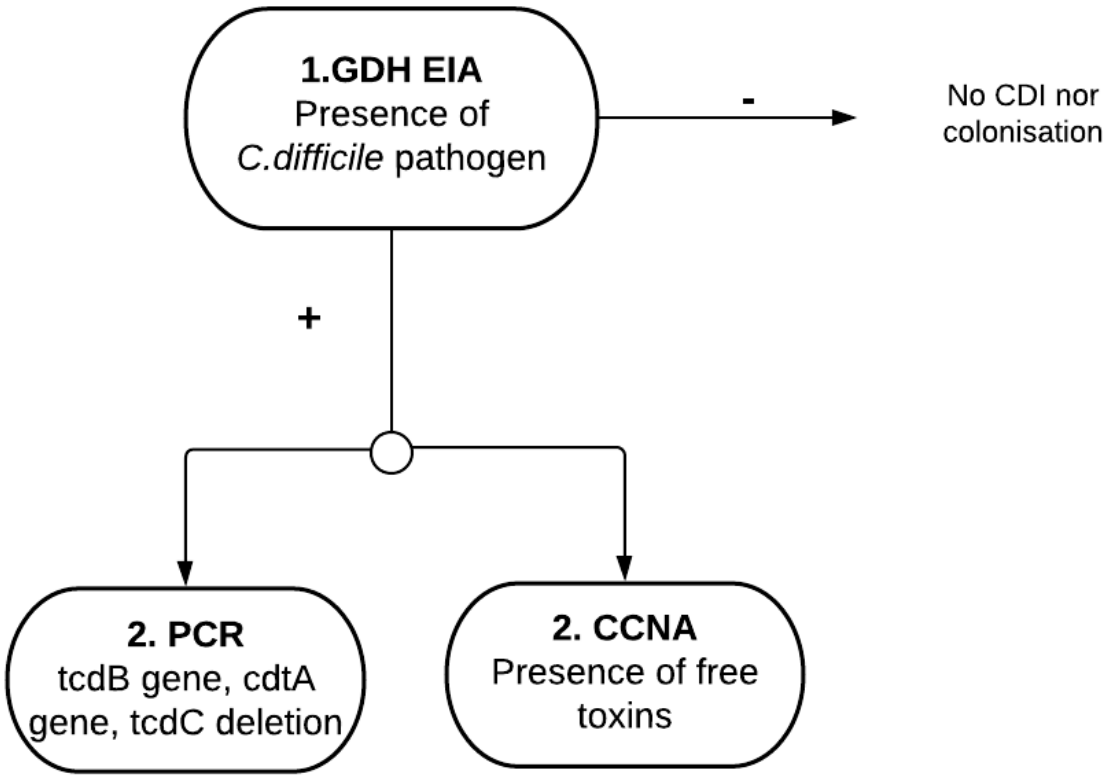



**Figure 5-3 Two-step testing algorithm run at LTHT to diagnose patients suspected with CDI**

**Table 5-2 Possible diagnoses for samples positive for GDH then tested with PCR and CCNA, outlining interpretation and clinical recommendations according to LTHT clinical guidelines (281, 282)**

| | Combined result | Interpretation | Clinical recommendation |
|---|---|---|---|
| PCR +ve/ CCNA +ve | • Toxigenic strain<br>• Free toxins detected | CDI positive | Contact and isolation measures are continued for the duration of inpatient stay |
| PCR -ve/ CCNA -ve | • Non-toxigenic strain<br>• No free toxins detected | CDI negative | Pending results on other GI pathogens, contact measures and isolation can be discontinued |
| PCR -ve/ CCNA +ve | • Free toxins detected | CDI positive | Contact and isolation measures are continued for the duration of inpatient stay |
| PCR +ve/ CCNA -ve | • Toxigenic strain<br>• No free toxins detected | Potentially colonised with *C. difficile* pathogen | Contact and isolation measures are continued for the duration of inpatient stay |

#### 5.3.3.4.2 Hypothetical test for CDI

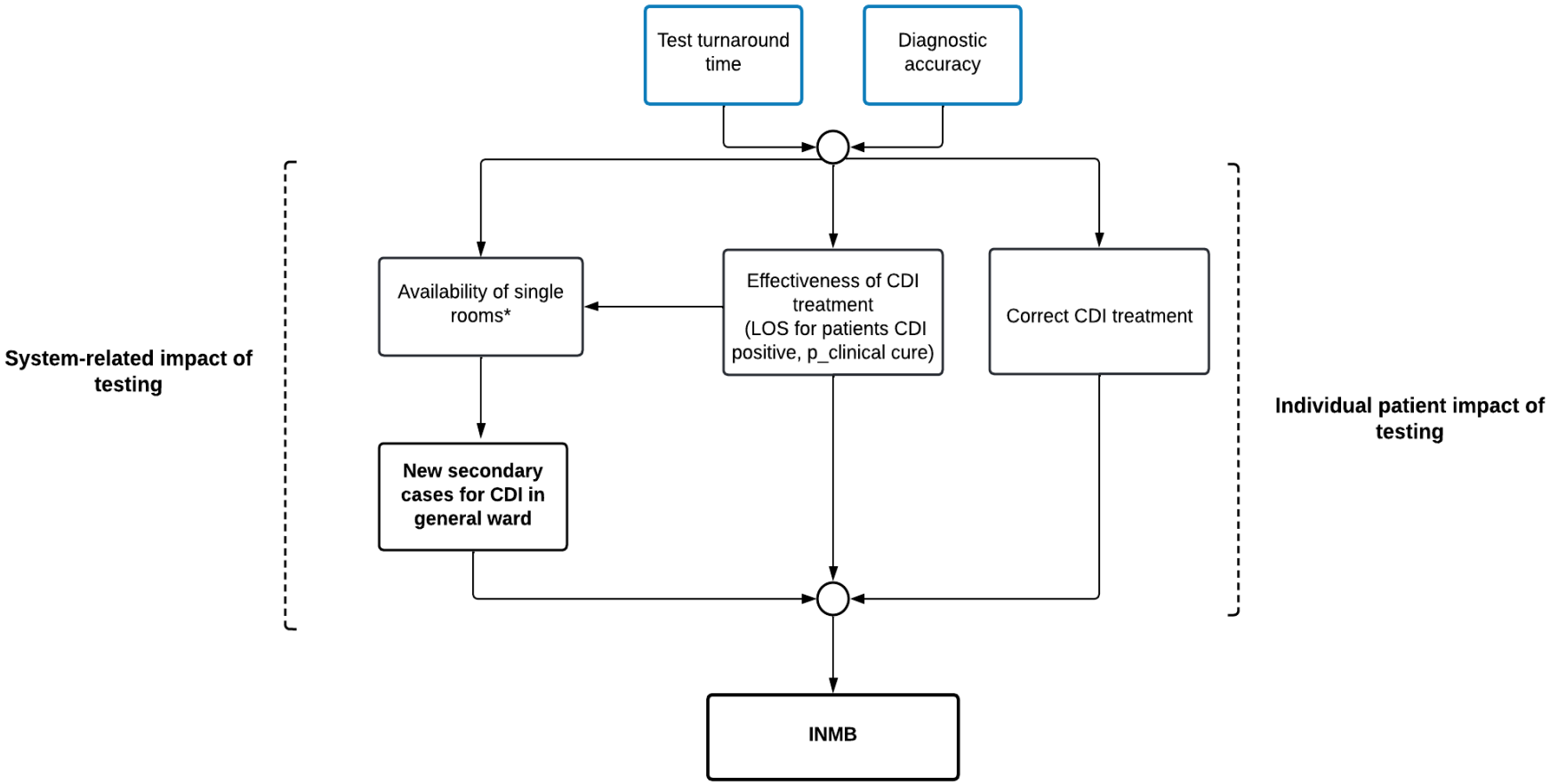
Based on the findings of the online survey presented in Chapter 4, a perceived need for more accurate and rapid new diagnostic tests for CDI emerged. In addition to this, a new POCT for CDI is currently under development as part of a funded research programme grant (as discussed in Chapter 1).

The intervention being evaluated was therefore a ward-based hypothetical POCT (henceforth 'HT') detecting toxins in stools. There are opposing views on which target diagnostic tests for CDI should detect, be it either free toxins in stool, the organism or genes encoding-toxins (215). This was also reflected in the findings of the online survey discussed in Chapter 4. Exploring the downstream benefits of detecting alternative targets, however, is outside the scope of this early economic model. It was therefore decided to model HT as a toxin detecting POCT since toxin detection has been reported to better correlate with disease severity and patient health outcomes (197, 218), as discussed in Chapter 3.

The focus of the model was on capturing the impact of HT on: (i) improving short-term clinical outcomes for patients via an expedited (and appropriate) administration of antibiotic treatment for CDI and (ii) supporting fast and appropriate escalation of infection control measures, to minimise in-hospital transmission and enable rapid de-isolation of non-infected patients. The model simulates two important mechanisms via which the HT may impact on clinical utility and cost-effectiveness outcomes: reduced test turnaround time, and increased diagnostic accuracy (see Figure 5-4).

As HT is a hypothetical test, no information is available currently on its diagnostic accuracy, turnaround time or test price. Various scenarios and sensitivity analyses were run in the model varying test turnaround time and diagnostic accuracy, as outlined in the Chapter 6.

**Figure 5-4 Simplified schematic of how HT affects the hospital system and individual patient health outcomes via reduced test turnaround time and improved diagnostic accuracy**



**Early diagnosis**

Upon receipt of positive test result for CDI, patients start antibiotic treatment for CDI – unless treatment has started in advance (i.e. *empirical treatment*) because patients presented with severe symptoms. Receiving test results in a timely manner supports clinicians in administering early treatment for CDI which, in turn, may improve short-term health outcomes for patients.

A HT with reduced turnaround time therefore: (1) increases the probability of clinical cure at day 10 of treatment regimen; as well as (2) reducing the hospital LOS compared to testing strategies with a slower turnaround time (260). In the context of early treatment, via early diagnosis, patients who tested positive to CDI experience a worsening in their quality of life for a shorter duration, while also decreasing the number of treatment days required for patients to recover to full health compared to patients who receive confirmation of CDI at later time points. This also leads to cost savings compared to standard care due to reduced total treatment cost and bed cost per patient.

On a system-related level, a HT capable of reducing turnaround time compared to standard care shortens patients' stay in presumptive isolation whilst awaiting the confirmation of infectious diarrhoea, as well as reducing the LOS patients positive to CDI remain in single room isolation. This results in more single rooms being available for new patients suspected with CDI thus reducing the number of secondary CDI cases in the general ward (236).

The possible impact of such a test on improving patient survival and risk of disease recurrence via expedited delivery of antibiotics was not captured in this analysis, due to a current paucity of data in the literature as to the effect of early treatment in this context (260, 266). As further data on this emerges, future iterations of this early economic model could be expanded to include this element.

**Increased diagnostic accuracy**

A HT associated with higher diagnostic specificity is able to decrease the number of FP cases (i.e. patients who were given a positive result for CDI, although they did not have the disease), thus reducing the number of unnecessary antibiotic treatments for CDI. This minimises the loss in health-related quality of life due to receiving incorrect antibiotic treatment. In addition, a HT test with improved diagnostic specificity increases the availability of free single rooms for new

suspected or confirmed infectious patients. With more single rooms being available, fewer patients are placed in the general ward while waiting for test results, thus minimising infection spread within the general ward. A HT with improved diagnostic sensitivity compared to standard care would reduce the number of FN cases and missed antibiotic prescriptions for CDI. On an individual patient level, this minimises the risk of health-related quality of life loss due to missed treatment; on a system-related level, a HT with improved diagnostic sensitivity reduces the risk of FN patients being released into the general ward, which decreases the risk of nosocomial transmission of CDI.

#### 5.3.3.5 Clinical decision-making upon receipt of test results

Upon receipt of a patient's test results, clinicians can decide to: (i) continue isolation; (ii) de-escalate isolation measures and move the patient into the general ward; or (iii) discharge the patient if the patient has spent their assigned LOS in isolation[18]. Based on current LTHT clinical guidelines, Table 5-3 gives an overview of recommended clinical actions for patients suspected with CDI considering both clinical symptoms and test results as simulated within the model.

[18] If a patient remains longer in single room isolation than the assigned LOS due to slow time-to-diagnosis, during the next check to assess resolution of symptoms, the model has the capacity to assess if the time the patient has spent within the simulation exceeds the assigned individual LOS. If so, the patient is assumed to be discharged.

**Table 5-3 Clinical decision-making for patients suspected with CDI considering symptoms and test results**

| | Test results |
|---|---|
| *Symptoms persist* | • **CDI +ve** – patient remains in isolation and antimicrobial treatment for CDI, unless empirical treatment was already started.<br>• **CDI -ve / GI panel +ve** – patient remains in isolation. Set daily checks to assess resolution of symptoms.<br>• **CDI -ve / GI panel -ve** – patient is confirmed with non-infective diarrhoea, isolation is not required. The patient is therefore de-isolated and moved into the general ward for their remaining hospital LOS. |
| *Symptoms disappear* | • **CDI +ve** – patient remains in isolation for the remaining length of hospital stay. Complete CDI treatment regimen.<br>• **CDI -ve / GI panel +ve** – patient remains in isolation for 2 extra days, and subsequently moves into the general ward – unless the patient is de-isolated in advance if case single rooms are in full capacity.<br>• **CDI -ve / GI panel -ve** – patient is de-isolated and moved into the general ward for their remaining hospital LOS. |

#### 5.3.3.6 Antibiotic treatment for CDI

If a patient presents with severe symptoms of CDI (e.g. evidence of severe colitis, high temperature), clinicians are assumed to start empirical antibiotic treatment before placing the patient into presumptive isolation and prior to receipt of test results (283). For patients with mild or moderate symptoms (e.g. approximately 3-5 stools per day), antibiotic treatment for CDI (i.e. vancomycin) starts upon receipt of positive test results for CDI.

The antibiotic treatment regimen for CDI lasts up to 10 days. At day 10, clinicians check if the patient is symptomatic and, if so, treatment is continued for 4 additional days. Once the patient is no longer symptomatic at day 10, treatment for CDI is discontinued and patients are assumed to recover to full health without risk of disease recurrence. Patients with a FP test result receiving unnecessary antibiotic treatment for CDI are assumed to remain symptomatic at day 10 of treatment, as they are not treated for the real cause of diarrhoea (e.g. non-infectious condition – as per expert opinion). Since patients with a FP test result for CDI remain symptomatic, they are assumed to receive treatment for 4 additional days (as per current clinical practice at LTHT).

Figure 5-5 shows a timeline of key events patients experience within the clinical management for CDI as simulated within the model.

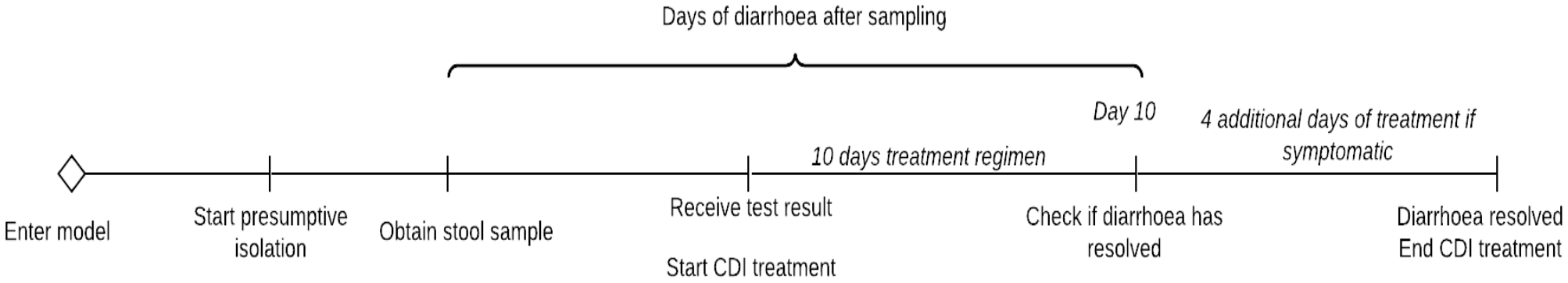



**Figure 5-5 Key events patients experience within the clinical pathway for CDI as simulated within the model**

#### 5.3.3.7 Dynamic bed allocation

The model has the capacity to resemble as much as possible the dynamic bed allocation processes underlying clinical practice, whilst capturing capacity constraints. The following subsections describe key events happening to patients placed either into single rooms, the general ward or cohort bays upon receipt of final diagnosis.

##### 5.3.3.7.1 Single room isolation

If single rooms reach full capacity, clinicians can decide to de-escalate infection-control measures for patients who are no longer infectious. Based on expert opinion, patients who are positive with other GI pathogens but asymptomatic can be assigned a lower priority from an infection-control perspective. Once single rooms are occupied, the model has the functionality to de-isolate any asymptomatic patient positive to other GI pathogens in order to release additional single rooms. This patient will then enter the general ward or be discharged depending on their remaining hospital LOS.

##### 5.3.3.7.2 Stay in general ward

If a patient within the general ward is confirmed with infectious diarrhoea, they may be transferred into a single room or cohort bay (depending on availability) to reduce risk of transmission. Alternatively, when single rooms and cohort bays are at full capacity, confirmed cases remain in the general ward until hospital discharge. The simulation then counts the total number of infectious CDI patients (i.e. TP and FN) within general ward over the model evaluation period (for more information see section 5.3.4).

If a patient within the general ward is reported as not infectious, they stay in the general ward until hospital discharge depending on their simulated LOS, without receiving antibiotic treatment for CDI.

##### 5.3.3.7.3 Cohort bay isolation

Upon confirmation of infectious diarrhoea, patients in the general ward can be transferred into one of the four cohort bays depending on current availability. Patients confirmed with CDI are grouped separately from patients positive to other GI pathogens, to reduce the risk of developing co-infections. Confirmed patients remain in cohort bays until hospital discharge.

### 5.3.4 Estimation of new secondary cases of CDI

As discussed in section 5.3.2, DES is not capable of directly tracking infection spread within the modelled system without intense computation. Nevertheless, DES is capable of efficiently approximating the infection spread. As such an approximation of the infection spread was captured in the model using a reproductive rate of infection spread ($R_0$) – which is represents an estimate of the number of secondary infected cases resulting from one primary infected patient (287).

Similar to existing decision models for CDI diagnostics (235, 264), at the end of the model simulation a simplified approach was taken to indirectly estimate the number of secondary infection cases resulting from infected patients entering the general ward, using $R_0$ (i.e. as opposed to adopting a SD model):

$$Secondary\ infection\ cases = primary\ infection\ cases\ in\ general\ ward\ \times R_0$$

The number of primary infection cases in the general ward comprises of patients positive to CDI who were placed in the general ward due to scarcity of single rooms and cohort bays, and those patients incorrectly de-isolated upon receipt of a FN test result.

The model estimated the potential number of secondary infection cases due to patients confirmed with CDI being in the general ward only, as patients positive to CDI within single rooms and cohort bays pose a minimal risk of wide-spread transmission. The rate of infection, however, is assumed to be constant regardless of the number of secondary cases occurring in the model. In addition, the model only counts the number of secondary cases at each run, and estimates the associated healthcare costs and QALY lost – without explicitly simulating the events and clinical processes occurring to new patients infected with CDI.

### 5.3.5 Model parameterisation

This section describes the assumptions and data used to populate the model. Table 5-4 provides an overview of model parameters and related sources. Key model parameters relating to patient characteristics and hospital configuration are based on the COMBACTE-CDI study datasets (presented in 5.3.5.1). The following subsections describe the other data sources used to parameterise the model (see 5.3.5.2-5.3.5.7).

#### 5.3.5.1 COMBACTE-CDI study datasets

COMBACTE-CDI was a multicentre European-wide epidemiological study assessing the impact of CDI on patients' health and clinical practice across 119 healthcare sites from 12 European countries, conducted from 2018 to 2021[19] (244). Full details on the COMBACTE-CDI trial are available in previous publications (244, 253). Access to COMBACTE-CDI data for the purpose of this thesis was granted by signing an assignment of intellectual property rights form to the University of Leeds (see Appendix L).

In particular, UK-individual patient data (IPD) from the COMBACTE-CDI case report form (CRF) dataset (n= 180 patients from 23 UK sites) was used to inform model parameters relating to patient characteristics (e.g. disease severity, symptoms duration, hospital LOS). This dataset contains anonymised individual patient data on: (i) demographics; (ii) admission history (e.g. reason for hospital admission, previous treatments received); (iii) laboratory testing (e.g. information on testing for any GI pathogens in advance to CDI testing, CDI diagnosis); (iv) clinical information (e.g. comorbidities, disease severity for CDI, diarrhoea duration); (v) drug history; (vi) CDI treatment escalation; (vii) surgery and GI interventions; and (viii) health outcomes (e.g. hospital discharge, re-admission, death).

Parametric survival analysis was conducted to determine appropriate distributions for time-to-event data in the model (i.e. duration of symptoms for CDI positive and negative patients, and hospital LOS). For each of these variables, different parametric models (Normal, Lognormal, Exponential, Gamma and Weibull) were fitted to the UK-based COMBACTE CDI IPD data using the R 'fitdistrplus' package (288). Based on a maximum likelihood estimation (MLE)

[19] Samples were collected only between 2018 and 2019.

process, the optimal fitting distribution for each variable was chosen based on: (i) an analysis of the Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC) metrics (the distribution with the lowest AIC and BIC values indicating the parameterisation with the optimal statistical fit to the data); (ii) an analysis of key summary statistics produced from the various parameterisations (i.e. mean, median, inter-quartile range, standard deviation [SD], and maximum values); and (iii) visual inspection of the goodness of fit of the different parametric models against the observed patient data (288). For more information see Appendix N.

In addition, UK-specific summary data from a European-wide survey part of COMBACTE-CDI study was used to inform key model parameters related to hospital configuration (e.g. numbers of samples run and single rooms available). This survey was sent out to community and hospital sites (n=158) across 12 European countries to assess current clinical practices for CDI patients and CDI costs (289). This dataset contains anonymised information on: (i) participating site information (e.g. type of healthcare facility, number of samples tested annually, bed capacity); (ii) disease detection, notification and management (e.g. adherence to clinical guidelines, reasons to exclude certain patient subgroups from testing for CDI, adherence to contact precautions); (iii) testing procedures (e.g. which tests for CDI are available, time to receive sample to the laboratory); (iv) surveillance practices (e.g. participation in international and/or national CDI surveillance schemes); (v) treatment policies and procedures (e.g. adherence to clinical guidelines, treatment options sorted by disease severity, administration of empirical treatment); and (vi) associated costs (e.g. average cost per single room isolation, testing costs, infection-control costs).

| Parameter | Estimate (SD) | Distribution type | Data source |
|---|---|---|---|
| *Disease-related parameters* | | | |
| CDI disease prevalence | 10.342% | Probability profile[20] | (265) |
| Other GI pathogens prevalence | 13.801% | Probability profile | (265) |
| CDI patients with mild symptoms | 39% | Probability profile | COMBACTE-CDI CRF dataset |
| CDI patients with moderate symptoms | 33% | Probability profile | COMBACTE-CDI CRF dataset |
| CDI patients with severe symptoms | 27% | Probability profile | COMBACTE-CDI CRF dataset |
| Length of stay CDI negative patients (days) | 28.6 (38.51) | Weibull | COMBACTE-CDI CRF dataset |
| Duration diarrhoea after sampling CDI negative patients (days) | 7.041 (11.98) | Weibull | COMBACTE-CDI CRF dataset |
| Duration diarrhoea after sampling CDI positive patients (days) | 12.27 (26.04) | Weibull | COMBACTE-CDI CRF dataset |
| UK-England utility weight general population aged 55-64 | 0.819 | Fixed | (291) |
| UK utility weight adult hospitalised patient with first episode of CDI | 0.42 | Fixed | (195) |
| Decrement utility weight due to inappropriate antibiotic treatment | 0.10 | Fixed | Assumption |
| Duration CDI disease | 10-14 days | Fixed | (265, 292) |
| Duration CDI clinical management | 30 days | Fixed | (260, 266) |
| Reproductive ratio for CDI (median) | 1.04 | Fixed | (293) |
| Probability of clinical cure – slow diagnosis | 85.3% | Probability profile | (260) |
| Probability of clinical cure – average diagnosis | 90.7% | Probability profile | (260) |
| Probability of clinical cure – rapid diagnosis | 95.6% | Probability profile | (260) |
| Length of stay CDI positive patients – slow diagnosis (mean days) | 30.3 (36.3) | Weibull | (260) |
| Length of stay CDI positive patients – average diagnosis (mean days) | 26.9 (28.9) | Weibull | (260) |
| Length of stay CDI positive patients – rapid diagnosis (mean days) | 23.2 (25.4) | Weibull | (260) |
| *Diagnostic accuracy* | | | |
| GDH EIA sensitivity | 94% | Fixed | (222) |
| GDH EIA specificity | 94% | Fixed | (222) |
| PCR sensitivity | 95% | Fixed | (222) |
| PCR specificity | 98% | Fixed | (222) |
| CCNA sensitivity | 86.4% | Fixed | (231, 294) |
| CCNA specificity | 99.2% | Fixed | (231, 294) |
| Multiplex GI panel testing sensitivity | 100% | Fixed | Assumption |
| Multiplex GI panel testing specificity | 100% | Fixed | Assumption |
| *Testing workflow* | | | |
| Monthly median stool samples tested in UK Teaching Hospitals (n) | 1430.5 (689) | Fixed | COMBACTE-CDI survey |
| Proportion of samples tested for CDI only | 5%-69% | Fixed | COMBACTE-CDI survey |
| Time to obtain stool sample | 0.5 (0-2 day) | Triangular | (266) |
| Time to transport sample to the laboratory | 15 min | Fixed | Assumption |
| Techlab C.diff Check GDH EIA operating time | 60 minutes | Fixed | (295) |
| Xpert *C. difficile* BT Cepheid PCR operating time | 43 minutes | Fixed | (296) |
| CCNA operating time (day) | 1.5 (1-2) | Triangular | Expert opinion |
| Multiplex GI pathogens panel operating time | 43 minutes | Fixed | Equal to PCR |
| Time for preparing sample batch | 30 minutes | Normal | Expert opinion |
| Xpert *C. difficile* BT Cepheid PCR operating time | 1 minute | Fixed | (297) |
| Time for preparing CCNA | 20 minutes | Fixed | Expert opinion |

| | | | |
|---|---|---|---|
| Time to review positive test results | 30 minutes | Fixed | Expert opinion |
| Time to review negative test results | 0 minutes | Fixed | Expert opinion |
| *Costs* | | | |
| Cost of bed day in adult isolation | £692.83 | Fixed | (298, 299) |
| Cost of bed day in general ward | £583 | Fixed | (298) |
| GDH EIA cost per kit | £4.84 | Fixed | (266) |
| PCR cost per run | £26.90 | Fixed | (266) |
| CCNA cost | £3.74 | Fixed | (300) |
| Multiplex GI panel cost per sample | £43.03 | Fixed | (266) |
| Vancomycin 125mg Drug Tariff unit price | £132.49 | Fixed | (301) |
| Additional cost per secondary case per day | £957.18 | Fixed | (194) |
| *Hospital configuration* | | | |
| Side rooms in typical UK Teaching Hospital (n) | 93 | N/A | COMBACTE-CDI survey |

**Table 5-4 Model parameters, related data sources and distribution type**

#### 5.3.5.2 Patient characteristics

Patients entering the model are grouped depending on their infection status for CDI and other GI pathogens. CDI disease prevalence is set at 10.342%, whereas prevalence for other GI pathogens is equal to 13.801%[21] (265). Both estimates were based on a clinical effectiveness meta-analysis review conducted in 2015 alongside a UK-based economic model on testing strategies for patients suspected with gastroenteritis (265). This meta-analysis review pooled data from two large USA-based studies evaluating the performance of multiplex GI panels (n=2,963) (302, 303).

It was assumed that there was no risk of co-infection between CDI and other GI pathogens. If a patient is confirmed to have CDI, they were assumed to be negative to other GI pathogens; conversely, if a patient is positive to other GI pathogens, they were assumed to be negative to CDI. Based on UK-specific COMBACTE-CDI CRF dataset, the expected prevalence of co-infection is negligible as only small proportion of patients negative to CDI tested positive to other GI pathogens (n=12, 9%); similarly, few confirmed CDI patients appeared to be positive to other GI pathogens (n=4, 13%). A similar modelling approach was taken by past models on CDI diagnostics which accounted for testing for other GI pathogens (265, 266).

[20] A probability profile in SIMUL8 is defined as "a type of distribution that sets the probability (the percentage change) of a value being sampled from a distribution" (290). For example, a probability profile of 10.342% was assigned to the disease prevalence of CDI. This means that, within each model run, there is a probability of 10.342% that each patient entering the hospital has CDI, as opposed to a probability of 89.658% of not having CDI.

[21] Other pathogens included: adenovirus, Campylobacter, Cryptosporidium, E.coli, Giardia, norovirus, rotavirus, salmonella and shigella (265)

Patients with positive result for CDI are further grouped depending on their disease severity. European estimates from the COMBACTE-CDI CRF dataset informed the proportions of patients with mild, moderate and severe CDI. European estimates were selected due to a paucity of UK data on disease severity.

#### 5.3.5.3 Time-to-event variables: duration of symptoms and length of stay

**Patients negative to CDI**

Duration of diarrhoea and length of hospital stay for truly CDI negative patients (i.e. TN and FP cases) was informed by UK-specific IPD from the COMBACTE-CDI CRF dataset (n=180). Truly CDI negative patients were reported to remain symptomatic on average up to 7.041 days after sampling (11.98 SD) (n=53) and to stay at the hospital around 28.6 days (38.51 SD) (n=68). Based on the parametric survival analysis, a Weibull distribution was selected as the best fitting distribution for the duration of symptoms and LOS for patients negative to CDI parameters (see Appendix N for full details).

**Patients positive to CDI**

UK-specific IPD from COMBACTE-CDI CRF dataset also informed the duration of symptoms for truly CDI positive patients (i.e. TP and FN cases), which was estimated at 12.27 days on average (24.04 SD) (n=12). Based on the parametric survival analysis, a Weibull distribution was applied to the duration of symptoms for patients with a positive CDI result parameter (see Appendix N).

LOS for truly positive CDI patients was based on a prospective time-series study conducted in a 750-bed hospital in France (n= 126 patients positive to CDI), which evaluated the impact of three testing regimens on the clinical management and short-term outcomes for patients with CDI. This study found that receiving rapid diagnosis for CDI (approximately 15 hours) had a positive impact on patient management while reducing hospital LOS and increasing the probability of being clinically cured at the end of the treatment regimen for CDI (260).

Depending on how quickly patients who tested positive to CDI (i.e. TP cases) receive a final diagnosis, a certain hospital LOS and probability of clinical cure at day 10 are assumed with rapid testing strategies leading to better short-term clinical outcomes compared to slower testing options. Figure 5-6 reports the different probabilities of clinical cure at day 10 of treatment regimen for CDI and

hospital LOS depending on the time to receive final diagnosis (260). Time cut-off points (e.g. 0.5 day, 1.2 day) reflect the mean time-for-return of results across the different testing strategies as reported in the prospective time-series study (260). It was assumed that patients with a FN result for CDI experienced a LOS equal to the LOS patients with a positive CDI result might experience in the context of slow diagnosis. This helped to capture the impact of missed diagnosis on patients' hospital LOS.

Based on the summary data from the prospective time-series study, different parameterisations were explored for these variables (Normal, Lognormal, Exponential, Gamma and Weibull). The best fitting distribution was selected based on an analysis of key summary statistics (mean, median, IQR, SD and maximum value). Based on this analysis, a Weibull distribution was applied to the LOS variables for patients positive to CDI (see Appendix N).

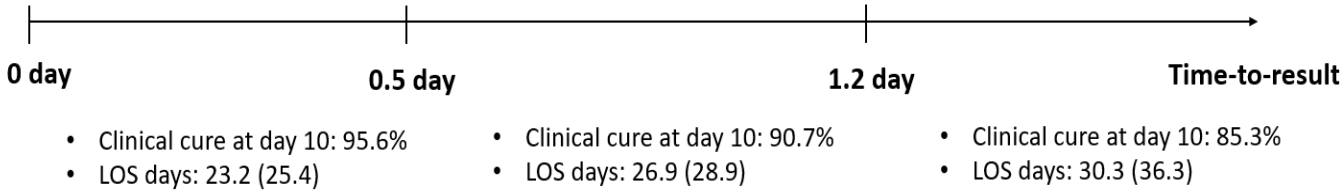



**Figure 5-6 Probabilities of clinical cure at day 10 of treatment regimen for CDI and hospital length of stay (SD), sorted by time-to-final diagnosis. Source: Barbut, F. et al. (260).**

#### 5.3.5.4 Hospital configuration

UK-specific summary data from the COMBACTE-CDI survey informed the total number of single rooms available (Table 5-5). In the UK, 8 Teaching Hospitals took part in this survey which provides the most comprehensive and up to date information on bed capacity and availability of single rooms within the UK. Estimates from this dataset were also validated with clinical experts.

According to the COMBACTE-CDI survey data, UK Teaching Hospitals had an average capacity of 279 side rooms. Not all of these single rooms would be used for patients suffering with GI symptoms however - some would be used for patients with other suspected or confirmed infectious conditions (e.g. acute meningitis, severe acute respiratory syndrome) (284). Based on expert opinion, it was assumed that one-third of the reported single rooms in the COMBACTE-

CDI survey would be allocated for patients with suspected diarrhoea. Based on this estimate, this model simulated 93 single rooms where patients suspected with infectious diarrhoea could be placed[22].

**Table 5-5 Single rooms availability in UK Teaching Hospitals (n=8). Source: COMBACTE-CDI survey**

| | UK Teaching Hospitals | Source |
|---|---|---|
| Mean (n) | 278.7 | COMBACTE-CDI survey |
| Standard Deviation | 93.8 | |
| Upper bound 95% CI (n) | 365.5 | |
| Lower bound 95% CI (n) | 192 | |
| Proportion of single rooms allocated for patients suspected with infectious diarrhoea | 33% | Expert opinion |
| Single rooms simulated (n) | 93 | Estimated |

#### 5.3.5.5 Inter-arrival time

The speed at which new patients enter the model was dictated by an assigned inter-arrival time. For the purpose of this model, the inter-arrival time should equal the expected frequency at which patients develop diarrhoeic symptoms within the hospital setting. This estimate was based on the average number of stool samples tested in the laboratory over a month from the COMBACTE-CDI (see Table 5-6). The observations from this data were highly right-skewed, with the mean being greater than the median due to a small number of high outliers in the data (n=2). Based on consultations with a clinical expert, the median number of samples being tested (for any condition) was considered to better represent the workload within a typical UK Teaching Hospital. This estimate (n= 1430) was thus applied in the model.

[22] The median estimate of single rooms was applied to represent the availability of single rooms for a typical UK Teaching Hospital while reducing the variability within the results due to outlier observations. This model did not attempt to capture the uncertainty variability around the number of available single rooms across the participating sites to the COMBACTE-CDI survey.

**Table 5-6 Monthly average stool samples being tested in UK Teaching Hospitals (n=8). Source: COMBACTE CDI survey**

| | UK Teaching Hospitals | Source |
|---|---|---|
| Mean | 1923 | COMBACTE-CDI survey |
| Standard Error | 688.7 | |
| Median | 1430.5 | |
| Standard Deviation | 1948.1 | |
| Minimum | 70 | |
| Maximum | 6239 | |
| Upper CI (95%) | 3551 | |
| Lower CI (95%) | 294 | |

This data, however, included stool samples being tested for other conditions, not only CDI. In addition to the above data, the COMBACTE-CDI survey included a question about the proportion of stool samples being tested only for CDI (see Table 5-6). Respondents were asked to select from a multiple-choice question about the proportion of samples being tested for CDI only on a monthly basis (including: 'less than 30% of samples', 'between 30-49%', 'between 50-69%', 'between 70-99% or '100%' of samples'). Because of the wording of the question, the frequency of the specific values within each range was unknown. As such, within each range, a uniform distribution was applied so that each value within the range was assumed to be equally likely to occur. In addition, for the first range (e.g. 0-30%) a minimum proportion of samples being tested for CDI was set equal to 5% to ensure functionality of the simulation. Out of 8 participating UK sites in the COMBACTE-CDI survey, the majority reported testing between 30%-49% samples for CDI (50%, n=4) or, alternatively, less than 30% of samples (25%, n=2). The remaining UK sites would test between 50%-69% of samples for CDI only (25%, n=2). This information was combined with the median number of samples reported in Table 5-6 to calculate the number of monthly samples tested for CDI in the model.

The monthly number of stool samples tested for CDI only was approximately equal to 539. As the monthly demand for stool testing was assumed to be

constant and independent from seasonality or other external factors, an exponential distribution was selected to calculate the inter-arrival time. In the model, the monthly average of stool samples tested for CDI only equates a mean inter-arrival time of 35 minutes (meaning that a new symptomatic patient enters the model on average every 35 minutes), assuming an exponential distribution. Full details on the calculations for these inter-arrival times are included in Appendix N.

#### 5.3.5.6 Features of the LTHT testing pathway for patients suspected with CDI

Diagnostic accuracies for the two-step LTHT testing algorithm (outlined in section 5.3.3.4.1) were based on published estimates taking TC as the diagnostic reference method (222, 231, 294). Due to a lack of data on sequential testing, independence between sequential tests was assumed.

Time to prepare a batch of samples to be tested and to set CCNA was based on expert opinion. A Normal distribution was applied to the 'time for preparing sample batch' variable to account for ± 25% variation from the mean, whereas a triangular distribution was applied to the operating time for CCNA based on the upper and lower bounds provided by the clinical expert. Manufacturers' instructions for use (295-297) and published literature (222, 304) informed the time to set up and processing time for each testing option within the LTHT testing algorithm.

Time to review CDI test results was based on expert opinion. If GDH results are negative, the LIMS sends an automatic alert to the clinician in the ward. Similarly, in case of PCR negative results, the LIMS automatically alerts the clinician. For positive results for CDI (e.g. GDH, PCR, CCNA), a biomedical scientist will review the results within 30 minutes – as per expert opinion.

#### 5.3.5.7 Multiplex GI testing panel

The multiplex GI panel, which detects the presence of other GI pathogens in patients with negative results for CDI, was assumed to have perfected diagnostic accuracy. This simplifying assumption was applied as the focus of this model lied in evaluating testing strategies for CDI, rather than other GI pathogens. A similar modelling approach was applied in the single economic model from the rapid literature review of CDI decision models review which evaluated the cost-effectiveness of a multiplex GI panel (266).

Based on expert opinion, one sample per time can be processed by the multiplex GI panel on an average time of 43 minutes. Three multiplex GI panels are available at LTHT to test samples for other GI pathogens – as per expert opinion – hence this number of devices were included in the model.

#### 5.3.5.8 New secondary cases of CDI

A reproductive rate of infection spread of 1.04 was applied within the model as a baseline value, whereas the minimum (1.99) and maximum (0.52) values of this parameter were varied in the univariate sensitivity analysis outlined in Chapter 6. This estimate was obtained from an epidemiological model on CDI in-ward transmission, which was based on a retrospective cohort study of an outbreak of CDI in a USA-based tertiary hospital in 2008 (293). Existing decision models for CDI diagnostics have all applied this same reproductive rate of infection spread from this epidemiological model (235, 264, 267). No equivalent data in the UK context could be identified.

#### 5.3.5.9 Health-related utility weights

Health-related utility weights were based on published estimates. Baseline utility weights for UK-England general population were obtained from the European Quality of Life-5 Dimensions (EQ-5D) index population norm data (291). Health-related utility weights associated with CDI were based on a prospective patient self-assessment of quality of life study conducted with UK adult inpatients with a first CDI episode (n=30) (195).

To account for the impact of receiving unnecessary antibiotic treatment for CDI and the marginally worse quality of life upon receipt of a FP diagnosis, a 10% reduction in quality of life until discharge compared to TN cases was applied. This estimate was based on expert opinion due to a paucity of data.

#### 5.3.5.10 Costing estimates

The modelled costs were based on published data and included testing costs, CDI treatment costs, bed costs and additional costs due to secondary infections. Costs were inflated to 2021 prices using the Bank of England inflator where appropriate (305).

*Testing costs*

Testing costs were based on published estimates from the Newcastle Microbiology Laboratory (266), inflated to 2021 from 2017 prices. Testing costs included the cost per testing kit and reagents only (266). Labour and transportation costs were excluded for the following reasons: (i) transportation costs were assumed to be negligible as the model simulated a UK Teaching Hospital with the availability of an on-site laboratory for testing (hence no transport would be required in this scenario); and (ii) labour costs associated with running a POCT and a laboratory-based testing strategy were assumed to be equivalent. A similar modelling approach was taken by the NICE-commissioned early economic model of COVID-19 POCTs (164).

*CDI treatment costs*

Costs for CDI treatment regimens were taken from the British National Formulary (BNF). Patients confirmed with CDI receive oral vancomycin 125mg every 6 hours for 10 days according to the LTHT clinical practice (283). Every vancomycin pack contained 28 capsules. Assuming one patient receive 4 capsules per day (24h/6h), one pack with 28 capsules is expected to cover 7 days of treatment per patient. For each patient, the total number of CDI treatment days was divided by 7 to estimate how many vancomycin packs were needed. This estimate was then multiplied by the BNF vancomycin 125mg Drug Tariff price equal to £132.49 (301). A similar costing approach was taken by a UK-based decision model on CDI diagnostics (266).

*Bed Costs*

Costs for bed-days in the general ward and isolation rooms were taken from UK-based decision model of CDI diagnostics (266), based on the NHS Reference Costs 2015/2016 (298). This estimate was based on the average costs for patients with GI infections without interventions, taken from the elective inpatient spreadsheet. To estimate the cost per adult bed day in the isolation ward, an

additional cost was added to the cost for bed days in the general ward to account for single room isolation (299).

*Additional costs due to secondary infections*

A UK retrospective cohort study estimated the cost per day for first CDI episode equal to £836 using micro-costing data from 45 hospitalised patients (194). This cost per day was multiplied by the typical duration of clinical management for a CDI patient (e.g. 30 days) to calculate the cost of a secondary CDI infection in the model (266).

### 5.3.6 Model outputs

The clinical effectiveness of each testing strategy was measured in terms of two key outputs: (1) QALY gains; and (2) the number of secondary infections prevented. Since CDI is a transient disease, QALYs were estimated by first calculating the quality-adjusted life days (QALDs) lost due to CDI, and then converting this into QALYs lost (265, 306). The individual QALD lost is estimated as follow (306):

$$QALD_i\ lost = (u_n - u_{CDI}) \times t$$

where $i$ is the individual patient, $u_n$ is the baseline utility weight for UK general population , $u_{CDI}$ is the health-related utility weight associated with CDI (195) , and $t$ is the number of days.

Truly infected patients who test positive for CDI are assumed to experience worsened quality of life for the duration of CDI treatment, and then to recover without risk of disease recurrence. Patients receiving a FN test result, meanwhile, are not provided with the necessary CDI treatment and are assumed to experience lower quality of life for the duration of their whole hospital stay.

Truly healthy patients who test negative for CDI (i.e. TN cases) experience no loss in quality of life due to CDI. Patients receiving a FP test result are administered unnecessary antibiotic treatment and are therefore assumed to experience an associated quality of life decrement for the duration of their treatment regimen. Due to a paucity of data, individuals receiving unnecessary CDI treatment in the model were assumed to experience a worsened quality of life compared to TN cases (see section 5.3.3.6).

To account for the new secondary infections, QALDs lost due to secondary cases were calculated as:

$$QALDs\ lost\ due\ secondary\ infections = Secondary\ CDI\ infections\ \times [(u_n - u_{CDI}) \times t_{CDI}]$$

where $t_{CDI}$ is CDI disease duration which is set to 14 days (292). Total QALDs lost comprises: (1) the sum of individual QALDs lost; and (2) the QALDs lost due to secondary infections. Total QALDs lost are then converted into QALYs lost by dividing by 365 days.

Finally, to enable standard cost-effectiveness calculations to be undertaken, the total QALY lost values were converted to QALY gains using an arbitrary common baseline value (in this case, 50) as shown below:

$$QALY\ gains = 50 - total\ QALY\ lost$$

Cost-effectiveness outputs of the model were expressed in terms of INMB from the UK NHS perspective (i.e. including direct healthcare costs), comparing each HT against standard care, using the NICE WTP lower threshold per QALY of £20,000. Clinical and cost-effectiveness outputs are calculated at the aggregate level – focusing on the total number of patients part of the *evaluation set* (i.e. those entering within the model entry period). The time horizon for the analysis is 7 months (see section 5.3.7.1 for further detail) and therefore a discount rate of 0% was applied.

Details on the evaluation process applied to derive minimum performance specifications for the HT based on the model outputs are provided in Chapter 6.

### 5.3.7 Model implementation

The model was implemented in the commercial software SIMUL8 version 28.0.0.4060 Student edition, and run on a VivoBook ASUS laptop X571GT_F571GT with a 2.60 GDHz Intel ® Core ™ i7 processor and 16 GB of memory under Microsoft Windows 10 Home (build 19,043). Total model run time was 20 minutes for HT testing strategy and 40 minutes for LTHT testing strategy. The model runs from 9am to 7pm (10 hours = 600 minutes in SIMUL8) for 5 days each week (Monday to Friday). Simulated time within the model progresses according to the *next time-to-event progression* mechanism.

The starting conditions of the model are first outlined (see section 5.3.7.1), followed by an overview of how the uncertainty was handled in the model (section 5.3.7.2).

#### 5.3.7.1 Model initialisation

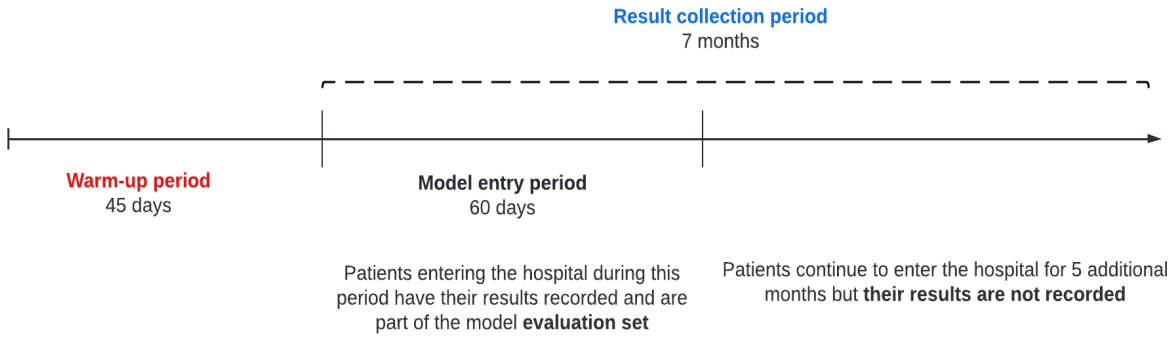
A *warm-up period* of 45 days (i.e. nine weeks) was applied in the model, to appropriately capture ongoing capacity constraints. Starting with the model 'cold' (i.e. all single rooms available and no patients in the system) would overestimate the hospital's capacity to test and isolate patients; simulating a warm-up period therefore helps to appropriately capture the steady-state capacity constraints patients face upon entering a busy hospital. In addition, all analyses were based on running 70 *model replications* (i.e. running the model 70 times, with each run using a different random number sequence). This approach ensures that the impact of *first-order* uncertainty on the model results is appropriately accounted for.

The warm-up period and replication number were set to values sufficient to provide stable outputs for the number of secondary infections within the general ward. As per good modelling practices for DES models (307), the warm-up period and model replication number were set to ensure a less than 1% difference observed between output values across increasing model replications.

Upon completion of the warm-up period, each model replication runs for a time horizon of 7 months (i.e. the *results collection period*). The model records outputs for every patient that enters the hospital during the *model entry period*, which is set equal to 60-days (two months) following the end of the warm-up period. Extending the results collection period to 7 months ensures that every patient within the model *evaluation set* (i.e. those entering within the model entry period) can have their full experience of the clinical pathway simulated. Whilst clinical management of CDI patients within the hospital usually lasts approximately up to one month (260, 266), a proportion of patients experience extended hospital stays (as reflected in the right-skewed tails of the LOS distributions applied in the model). A seven-month extended results collection period was therefore required to ensure outputs for patients entering during the model entry period were captured. Whilst additional patients were allowed to enter the model after the model entry period (to continue to simulate the busy hospital environment), only results for those patients who entered the model during the model entry period

were recorded. Figure 5-7 provides a schematic of the timeline for the model analysis. Details on the model analysis and estimation of required warm-up period and number of model replications are included in Appendix O.

**Figure 5-7 Relevant time points for model analysis**




### 5.3.7.2 Handling uncertainty

Although the model presented in section 5.3 is deterministic, *first-order* uncertainty was captured by running different iterations of the model 70 times – with each run using a different random number sequence. The example below highlights the importance of minimising *first-order* uncertainty of an event probability.

Patients entering the model are assigned a probability of having CDI which has a fixed mean value attributed to it (i.e. 10.342%). At the arrival of each new patient however, the DES modelling technique samples a random number from a uniform distribution (i.e. 'throws the dice'), compares that random number to the assigned mean probability (0.10342) and uses this to determine whether a certain individual in the model has CDI or not. For example, assuming a random number of 0.10, this value is below the probability of having CDI (0.10342) and therefore the patient will be classed as having CDI. The opposite is true if the random number exceeds the probability of having CDI. Although the overall likelihood that a patient has CDI is set of a fixed mean, there is a *first-order* uncertainty introduced at the individual patient level (i.e. the fact that SIMUL8 'throws the dice' at each patient's arrival), which in part determines whether or not they are truly with CDI or not. In no-constraints DES models (e.g. with no limit on number of single beds), simulating a large number of patients entering the hospital will estimate on average the true number of CDI positive patients which, in turn,

eliminates this issue. As such, in order to reduce the impact of this uncertainty within constrained-DES models, sufficient model iterations need to be run, using different random number sequences for each run of the model.

Since a PSA was not conducted as part of the model analysis, *second-order* uncertainty[23] was not accounted for. Conducting a PSA for each possible value of diagnostic accuracy and test turnaround time within the analysis would have significantly increased the computational burden and extended the model running time. Extensive sensitivity and scenario analyses were instead conducted to explore the uncertainty around key parameters and clinical assumptions and its impact on model outcomes (see Chapter 6 for more details).

[23] *Second-order* uncertainty is also referred as *parameter uncertainty* – defined as "the uncertainty in the estimation of the parameter of interest" (280). The impact of *parameter uncertainty* can be estimated by deterministic sensitivity analysis, or preferably, by running a PSA (280).

### 5.3.8 Model validation

The model structure and parameters underwent extensive validation, including:

- **face validity** – consultations with clinical experts ensured validity of the conceptual model and its input data;
- **internal validity** – the model was checked for any errors in coding by running extreme sets of parameters values (*extreme value testing*), replacing distributions with a constant number, and tracking patients through the model to assess whether the coding logic was correct *(testing of traces*). Model coding was also internally reviewed by an expert modeller within the Academic Unit of Health Economics, University of Leeds.
- **operational validity** – marginal and extreme changes to baseline parameters values were explored to assess: (i) how model parameters and structural assumptions affect the model outputs; and (ii) whether the model outputs would follow the expected trajectory following changes in parameters and structural assumptions. This validation exercise is described in the section 5.3.8.1 below.

#### 5.3.8.1 Operational validation

Figure 5-8 gives a simplified schematic of the mechanisms by which model parameters and structural scenarios have an impact on model outputs. The following sections describe how the availability of single rooms, new secondary cases and total QALY lost due to CDI vary depending on changes of model parameter or structural assumptions.

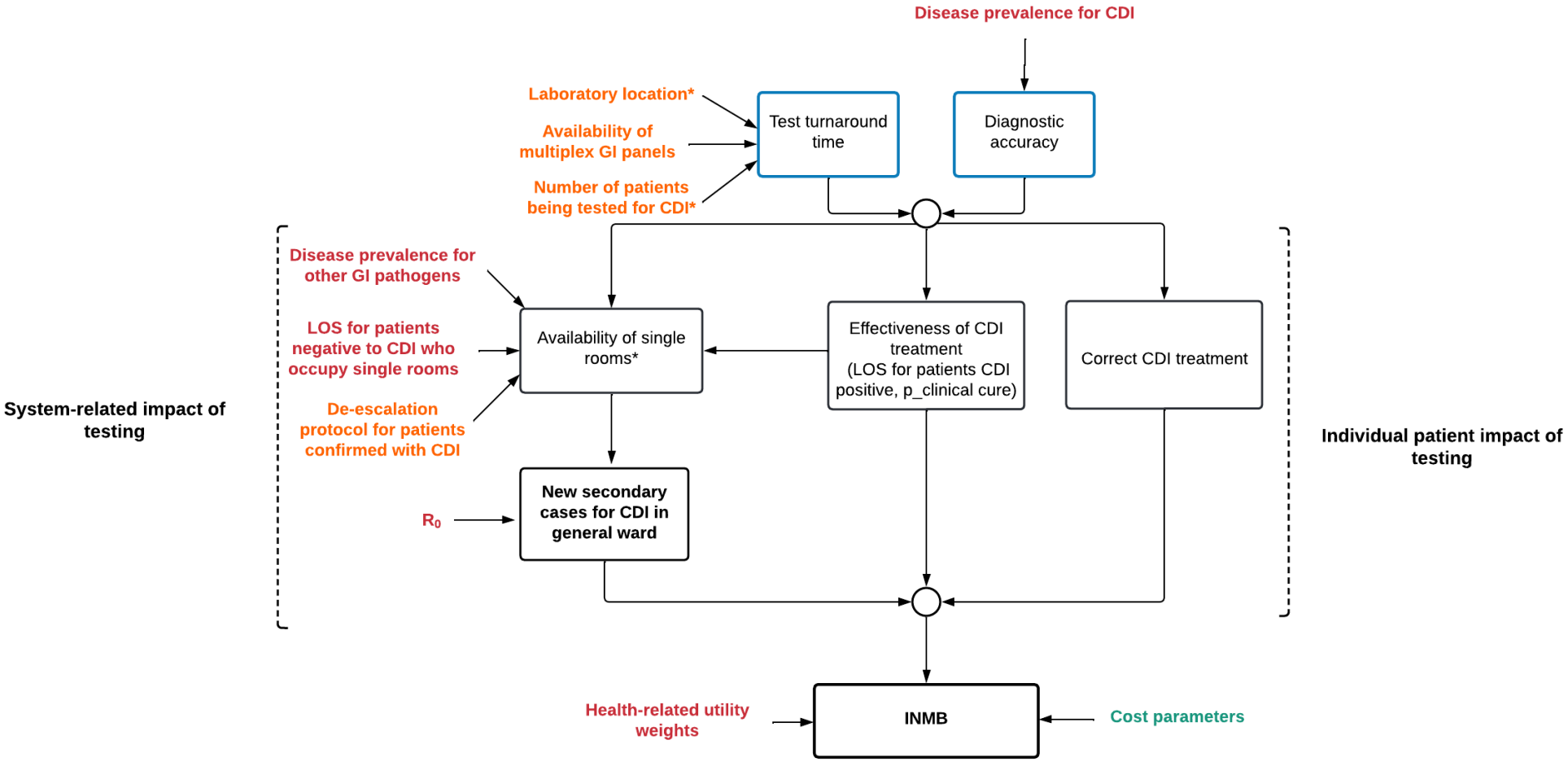



**Figure 5-8 Simplified schematic of the impact of model parameters and structural scenarios on clinical- and cost-effectiveness outputs. Disease- and testing-workflow-related parameters are associated with red and blue colourings, respectively. Sections highlighted in green and orange are associated with cost parameters and structural scenarios, respectively. Sections with an asterisk (*) are dependent on the size of the hospital being simulated**

#### 5.3.8.1.1 Availability of single rooms

In addition to the two mechanisms via which testing strategies are assumed to have an impact on the clinical workflow (i.e. turnaround time and diagnostic accuracy – discussed in section 5.3.3.4.2), the availability of single rooms is also dependent on the model parameters and structural assumptions, including:

- **disease prevalence for CDI** – in the context of higher disease prevalence, more confirmed CDI patients would require single room isolation until the end of their hospital LOS;
- **disease prevalence for other GI pathogens** – following a higher disease prevalence of other GI pathogens, more patients remain in single room isolation upon receipt of positive test result for infectious diarrhoea related to GI pathogens;
- **duration of symptoms for patients negative to CDI** – with longer duration of symptoms, patients confirmed with other GI pathogens occupy single rooms for longer;
- **slow test turnaround time** – the longer it takes for patients to receive test results for confirmation of infectious diarrhoea, the less rapidly patients are released from single room isolation. Slow test turnaround time can result from: (i) delays in obtaining the stool samples from patients; (ii) longer time to transport samples to the laboratory – especially in the context of hospitals using an off-site laboratory; (iii) more samples being tested for CDI and other GI pathogens; (iv) fewer multiplex GI panels available to rule out infectious diarrhoea due to other GI pathogens; and (v) extended operating times for each testing option;
- **de-escalation protocol for patients confirmed with CDI** – if clinicians decide to release patients positive to CDI from single room isolation at the end of the hospital LOS (as per LTHT practice), rather than after 48 hours from resolution of symptoms (i.e. as per UK clinical guidelines and the most common survey findings [see Chapter 4]), the availability of single rooms is reduced.

In the context of a reduced availability of single rooms, more secondary cases for CDI remain in the general ward.

#### 5.3.8.1.2 New secondary CDI cases

The reproductive ratio for CDI also has a direct impact on the number of new secondary cases for CDI in the general ward. An upper value for $R_0$ increases the number of confirmed cases due to one confirmed infected case in the general ward.

In the context of a shortage of single rooms, more patients confirmed with CDI remain in the general ward, thereby increasing the number of new secondary cases. This also increases the total healthcare costs and reduces the total QALY gains, leading to lower INMB for HT compared to LTHT testing strategy.

#### 5.3.8.1.3 Total QALY loss due to CDI

In addition to the impact of diagnostic accuracy and test turnaround time, fluctuations in the total individual QALYs lost due to CDI vary depending on changes in the following factors: (a) number of patients who truly have CDI; (b) number of FP patients; (c) number of FN patients; (d) LOS for patients positive to CDI; and (e) health-related utility weights for CDI. Several factors can increase the total individual QALYs lost, including:

- **disease prevalence for CDI** – in the context of a higher disease prevalence, more patients truly infected with CDI would experience a worsen quality of life due to having CDI;
- **health-related utility weights for CDI patients** – following a lower utility weight for CDI, patients confirmed with CDI experience a worsened quality of life compared to baseline. Similarly, having a higher decrement for unnecessary antibiotics for CDI leads FP patients to experience a worsened quality of life compared to a TN patient.

Upon an increase of total individual QALY lost, the total QALY gains for HT compared to LTHT testing strategy are lower thereby reducing the INMB associated with the HT.

## 5.4 Chapter summary

- This chapter presented the development of a *de novo* early economic model for a hypothetical rapid test for CDI.
- A rapid literature review of decision models evaluating the cost-effectiveness of diagnostic tests for CDI was first conducted to identify common modelling approaches and facilitate a critical appraisal of the existing models. Across the included studies (n=8), common methodological limitations were found, including: (i) a failure to capture the relevant events patients experience as part of the clinical pathway for CDI; and (ii) a weak evidence base in relation to the health-related utility weights for CDI. Furthermore, the majority of models failed to capture capacity considerations and key operational processes when evaluating the cost-effectiveness of CDI diagnostics;
- A validated *de novo* DES model was developed in SIMUL8 to map the flow of patients presenting with CDI symptoms through the hospital setting, accounting for the constrained availability of isolation rooms. Model parameters were informed by UK-specific data from COMBACTE-CDI study, literature sources, and expert opinion where necessary.
- The model compares a ward-based hypothetical POCT test detecting toxins in stools, against the two-step testing algorithm currently run at LTHT. For each testing strategy, the model recorded QALY gains and the number of secondary CDI infections in the general ward. Cost-effectiveness outputs of the model were expressed in terms of INMB.

In the following chapter (Chapter 6), the *de novo* model is analysed to facilitate the identifications of the minimum performance specifications for the HT, based on clinical and cost-effectiveness outputs. A detailed overview of the analysis approach and results is presented in Chapter 6.

# Chapter 6

# Early economic model for a new hypothetical rapid diagnostic test for CDI – analysis and results

## 6.1 Chapter outline

The previous chapter presented the structure, development, parameterisation and implementation of the early economic model evaluating the cost-effectiveness of a hypothetical rapid test for CDI compared to the standard care testing strategy at LTHT. This chapter outlines the model analysis (see section 6.2), which is divided into three phases: (1) a novel three-stage pragmatic framework to back-calculate minimum performance specifications for: (i) diagnostic sensitivity, (ii) diagnostic specificity, (iii) test turnaround-time, and (iv) maximum test cost based on the modelled clinical- and cost-effectiveness results (section 6.2.1); (2) sensitivity and scenario analyses on the clinical and cost-effectiveness outputs associated with the minimum performance specifications, to identify key drivers of impact in the model (section 6.2.2); and (3) sensitivity analyses exploring the impact of selected key drivers on the minimum performance specification results (section 6.2.3). Results of these analyses are presented in section 6.3. The final section of this chapter discusses the key findings and limitations of the model (section 6.4).

## 6.2 Model analysis

The goal of the analysis was to identify minimum acceptable performance specifications for the HT with respect to: (i) diagnostic sensitivity; (ii) diagnostic specificity; (iii) test turnaround time; and (iv) maximum test cost. Using the model, the minimum performance specifications were derived based on the simulated clinical- and cost-effectiveness outputs. Given the run-time for the model, exploring every possible combination of available values for these four variables would not be feasible nor useful. The analysis was designed to pragmatically and efficiently isolate the minimum acceptable performance specifications for each of the aforementioned outcomes of interest, and test the robustness of those results. The analysis therefore consists of three consecutive phases, including:

1. **Minimum performance specifications framework** – a novel three-stage framework (henceforth ‘MPS framework’) was developed to pragmatically

identify the expected minimum performance benchmarks for the HT based on deterministic clinical- and cost-effectiveness outputs. All other model parameters are held at their baseline value at this stage.

2. **identification of key model drivers** – deterministic univariate sensitivity and scenario analyses were conducted to assess the impact of varying each model parameter, and structural scenarios, on the clinical- and cost-effectiveness outputs of the model associated with the minimum performance specifications identified in phase 1. The aim of this phase was to identify key drivers of the model outputs which could potentially impact on the results from phase 1.
3. **sensitivity analyses on the minimum performance specifications** – sensitivity analyses were conducted to assess the impact of the top two influential model parameters and structural scenarios (identified in phase 2) on the minimum performance specifications determined in phase 1. Results from this phase indicate how the minimum performance requirements identified in phase 1 might change depending on different values of the model key drivers.

### 6.2.1 Phase 1: MPS framework

A novel three-stage pragmatic framework was developed and applied to identify the minimum performance requirements for the HT (compared to LTHT testing strategy), based on the incremental clinical and cost-effectiveness model outputs (see Figure 6-1). Generalisability of the MPS framework is discussed in Chapter 7.

The following sections describe each stage of the MPS framework in detail.

| Stage | Aim(s) | Methods |
|---|---|---|
| **1 - Assessment of clinical effectiveness outputs** | To exclude any region of diagnostic accuracy resulting in clinically inferior outputs | • Determine when HT is considered less clinically effective<br>• Two-way sensitivity analysis varying diagnostic sensitivity and specificity<br>• Hold third test performance dimension† constant to best-case value<br>• Assume zero testing cost for HT |
| **2 - Assessment of cost-effectiveness outputs** | To exclude any region of diagnostic accuracy not able to maintain cost-effectiveness | • Three-way sensitivity analysis varying diagnostic sensitivity, specificity and third test performance dimension*<br>• Assume zero testing cost for HT |
| **3 - Identifying minimum test specifications and maximum acceptable prices** | (1) To convert the results of the previous stages into '*minimum*' specifications<br>(2) To estimate the headroom and threshold prices using the WTP threshold | • *Minimum* - the lowest levels of diagnostic sensitivity and specificity which maintained clinical- (stage 1) and cost-effectiveness (stage 2)<br>• *Headroom price* - consider the perfect level of diagnostic accuracy - sorted by third test performance dimension†<br>• *Threshold price* - consider the lowest minimal diagnostic accuracy pair as identified in stage 1 and 2 - sorted by third test performance dimension† |

†In the context of the thesis model, the third test performance being simulated is test turnaround time
*The three-way sensitivity analysis is constrained to a pre-specified set of values for test turnaround time (15, 90 and 180 minutes)

**Figure 6-1 Three-stage MPS framework used to estimate minimum performance specifications based on cost-effectiveness outputs**

*Stage 1: Assessment of clinical effectiveness outputs*

HT was considered to be clinically inferior compared to standard care when: (i) HT led to negative incremental QALYs, and (ii) HT increased the number of secondary infections. Whilst a technology may be clinically less effective than standard care but nevertheless present a cost-effective alternative (e.g. if sufficient cost savings are accrued), the exclusion of clinically inferior options in this case was considered appropriate given that the HT is not expected to deliver substantial cost savings, and clinically inferior strategies would not be expected to be acceptable in this context – based on discussions with clinical stakeholders.

An initial analysis was run to rule out any combinations of diagnostic accuracy that resulted in the HT strategy being clinically inferior to the standard care testing strategy given a rapid test turnaround time. Diagnostic sensitivity and specificity pairs were varied over a pragmatic range, while assuming the quickest test turnaround time for detecting toxins in stools (i.e. 15 minutes – as per expert opinion). This allows identification of the widest region of diagnostic accuracy where the HT is expected to be more clinically effective than standard care, to take forward into the next stage of the analysis. At this stage, simulating a longer turnaround time for HT is expected to reduce the region of acceptable diagnostic accuracy where HT is more clinically effective due to the downstream negative impact of receiving a delayed CDI diagnosis – as discussed in Chapter 5.

*Stage 2: Assessment of cost-effectiveness outputs*

In this stage, test turnaround time and the identified acceptable combinations of diagnostic accuracy from Stage 1 were varied to assess their joint impact on the cost-effectiveness of HT, assuming a zero-test cost for HT (i.e. including CDI treatment costs, bed costs and costs related to secondary infections only). For each combination of diagnostic accuracy, the INMB of HT (compared to LTHT testing strategy) was estimated using the NICE WTP lower threshold (£20,000 per QALY gained). This was repeated for three different test turnaround time values: (i) 15 minutes (considered to be the fastest turnaround time for a POCT in this context, based on expert opinion); (ii) 90 minutes (as a mid-point value); and (iii) 180 minutes (equal to the current turnaround time for laboratory-based PCR).

At this stage, the aim was to rule out performance levels where the HT was not expected to be cost-effective, despite the test itself being costed at zero. This

identified combinations of diagnostic accuracy and test turnaround time to be excluded, as the decision maker or end-user would have to be paid to use HT in these scenarios (i.e. negative unit cost per test).

*Stage 3: Minimum acceptable test specifications and maximum acceptable prices*

For each value of test turnaround time evaluated, the performance specifications that resulted in HT being cost-effective were defined based on the results of Stage 1 and 2. The 'minimum' levels corresponded to the lowest values of sensitivity and specificity which maintained a positive clinical benefit (stage 1) and cost-effectiveness (stage 2).

In addition to this, the headroom and threshold unit prices were estimated using the NICE WTP threshold of £20,000 per QALY gained for each test turnaround time value. Headroom price represents the maximum price for an intervention to be cost-effective, assuming that the intervention meets perfect performance specifications (in this case, best-case values for diagnostic accuracy). Perfect conditions, however, are difficult and/or often impossible to attain. Based on the minimum performance specifications for diagnostic accuracy identified from stage 2, a maximum price for HT (i.e. *threshold price*) to be cost-effective was estimated for each test turnaround time value. This shows the maximum price manufacturers can expect to charge for a new test which meets the performance specifications for diagnostic accuracy, bounded from the lowest acceptable level of performance up to perfect performance, for a range of test turnaround time values.

### 6.2.2 Phase 2: Identification of key model drivers

Minimum performance requirements for new or hypothetical tests may vary depending on different values of model parameters or structural factors relating to the clinical context. It is not feasible, however, to re-run the MPS framework for every possible model parameter value and/or structural scenario. The aim of this phase, therefore, was to identify the key drivers of cost-effectiveness changes in the model, to inform which parameters and scenarios should be prioritised for sensitivity analyses around the primary outcomes in phase 3.

*Deterministic univariate sensitivity analysis*

Deterministic univariate sensitivity analyses were conducted to explore the impact of varying each of the model parameters on the clinical- and cost-effectiveness outputs. These analyses were run assuming a 15-minute test turnaround time, fixing the diagnostic sensitivity and specificity values to the minimum performance specifications identified in phase 1, and fixing the test price at the threshold value (see section 6.2.1).

For the remaining model parameters, lower and upper bounds for the parameter values were defined. Each sensitivity analysis was then run using 70 model replications (using different random number sequences), recording the incremental QALY gains, number of secondary cases, incremental costs, and INMB at £20,000 WTP per QALY gained. Full details on the lower and upper parameter values applied in this analysis are provided in Appendix O. For several parameters, lower and upper bounds were taken from a recent economic model on CDI diagnostics (266), or from the literature, where available. For the remaining parameters, the range of parameter values assessed was based on a +/-25% and +/-50% deviation from the baseline parameter value, separately.

*Scenario analyses*

The clinical care pathway for patients suspected with CDI varies across different hospitals, as discussed in Chapter 4. Based on consultations with clinical experts and results from Chapter 4, key structural variations in the care pathway for patients suspected with CDI were identified. Scenario analyses were therefore conducted in this phase to explore the impact of key structural differences on the modelled outputs. As before, these analyses were conducted assuming that sensitivity and specificity equal to the minimum performance specifications at 15-minutes test turnaround time as derived in phase 1, and fixing the test price at the threshold value (see section 6.2.1). Full details on the scenarios explored are provided in Appendix O. Scenarios investigated in this phase include:

**S.1 Alternative de-escalation protocol for patients confirmed with CDI** - at LTHT, patients confirmed with CDI remain in single room isolation until the end of their hospital stay. Based on the survey results (Chapter 4) however, clinicians in other hospitals de-isolate patients confirmed with CDI 48 hours after the resolution of symptoms. This difference in de-escalation practices might have an impact on the availability of single rooms.

**S.2** **Increased availability of multiplex GI panels** – the quantity of multiplex GI molecular platforms available may vary across different clinical settings depending on laboratory's financial constraints and number of stool samples being tested. This scenario explored the impact of having an additional multiplex GI molecular platform – resulting in a total of four multiplex GI molecular platforms being available to detect other GI pathogens in patients suspected with CDI.

**S.3** **Reduced availability of single rooms** – this scenario simulated the impact of HT in a smaller hospital with reduced availability of single rooms compared to LTHT. The baseline value of single rooms available (n=93) was reduced by 25%, resulting in 70 available single rooms in this scenario.

**S.4** **UK District Hospital** – compared to UK Teaching Hospitals, UK District Hospitals typically have: (i) fewer single rooms (93 single rooms within UK Teaching Hospital vs 64 single rooms within UK District Hospital); (ii) fewer patients being tested for CDI (1430 monthly average of stools being tested within UK Teaching Hospital vs 850 stools within UK District Hospital); and (iii) an off-site laboratory (see Appendix O for further details). This scenario analysis explored the impact of these differences.

**S.5** **Application of second-best distribution to time-to-event parameters from parametric survival analysis** – this scenario investigated the impact of applying a different parametric distribution to the duration of symptoms and LOS parameters. For each time-to-event parameter, the second-best fitting distribution was selected (see Chapter 5). Key properties of the time-to-result parameters are summarised in Appendix O.

**S.6** **Inclusion of labour costs for running laboratory testing** – this scenario investigated the impact of including additional labour costs for running laboratory-based testing strategies. Since HT is a ward-based POCT, additional costs are applied solely to the standard care testing strategy.

For each scenario analysis, 70 model replications were run recording the incremental QALY gains, number of secondary cases, incremental costs and INMB at £20,000 WTP per QALY gained.

### 6.2.3 Phase 3: Sensitivity analyses on the minimum performance specifications

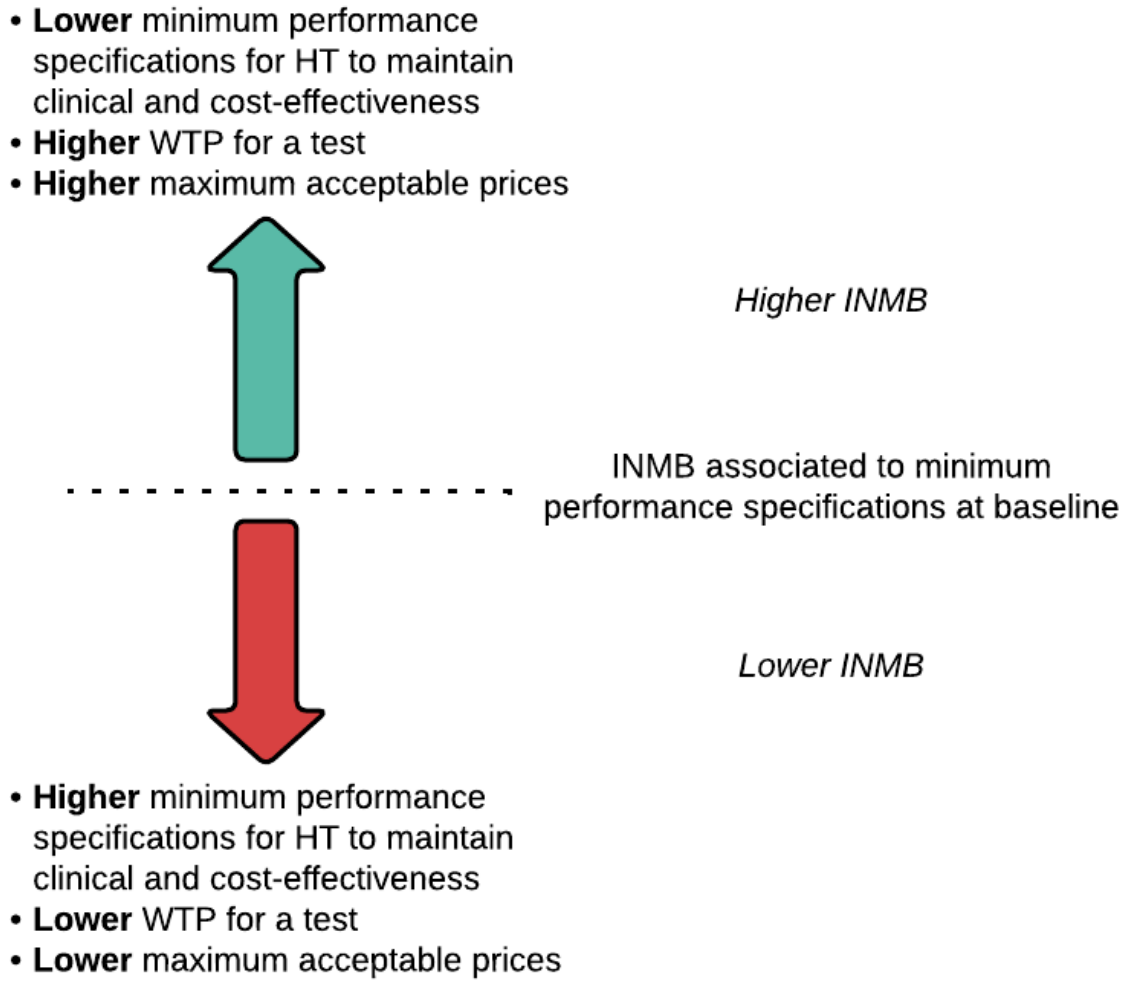
Based on the results of the previous phase, sensitivity analyses were run to assess the impact of top two most influential model parameters[24] (n=2) and structural scenarios (n=2) on the minimum performance specifications previously identified in phase 1.

For each of the selected influential model parameter values (e.g. lower and upper bounds) and scenarios, deterministic sensitivity and scenario analyses were run to assess how the minimum performance specifications (as identified in phase 1) would change in response to the change in parameter value/ scenario. Figure 6-2 presents how the findings of the sensitivity analyses at this phase could be used to narratively draw inferences on the expected trajectory of the minimum performance test specifications. In the context of parameter values and structural scenarios where the marginal benefit of HT was increased (i.e. higher INMB than baseline), based on the results of phase 2, the minimum performance specifications for HT to maintain clinical- and cost-effectiveness are expected to be lower compared to the baseline, while the HT can be priced to a higher maximum acceptable price (i.e. headroom and threshold prices) and still be considered cost-effective. Conversely, in the context of parameter values and structural scenarios which led to a reduced marginal benefit for HT (i.e. lower INMB than the baseline), higher minimum performance specifications and lower maximum acceptable prices are required for the HT to maintain clinical and cost-effectiveness, compared to the respective baseline values.

As before, this analysis was run assuming a 15-minute test turnaround time, and fixing the diagnostic sensitivity and specificity equal to the minimum performance specifications identified in phase 1, and holding the test price equal to the threshold value estimated in phase 1. For each change in parameter value/scenario, 70 model replications were run recording the incremental QALY gains, number of secondary cases, incremental costs and INMB at £20,000 WTP per QALY gained.

[24] The top two influential parameters were identified based on the results of the deterministic univariate sensitivity analysis using estimates from the literature and the +/-25% deviation from the baseline parameter value. The +/-25% range was selected as per common practice among economic evaluations.

**Figure 6-2 Simplified schematic of the narrative inference of the impact of changes in model parameter values/scenarios on minimum performance specifications for HT based on cost-effectiveness outputs**



## 6.3 Results

This section presents the results from each stage of the MPS framework (see section 6.3.1), followed by the results from identification of key model drivers (section 6.3.2), including the deterministic univariate sensitivity (section 6.3.2.1) and scenario analyses (section 6.3.2.2). Results of the sensitivity analyses on the minimum performance specifications are then presented in section 6.3.3.

Note that results are presented at the aggregate level – focusing on the total number of patients part of the *evaluation set* (i.e. those entering within the model entry period) – unless stated otherwise.

### 6.3.1 Phase 1: MPS framework

The following sections describe the results of each of the stages of the MPS framework separately.

#### 6.3.1.1 Stage 1 – Clinical outputs

At this stage of the analysis, HT was considered to be clinically inferior compared to standard care when: (i) HT led to negative incremental QALYs, and (ii) HT increased the number of secondary infections - as stated in section 6.2.1.

Table 6-1 shows the incremental clinical effectiveness for HT at each of the diagnostic accuracy combinations explored, assuming a 15-minute turnaround time. The lower the diagnostic specificity, the higher the minimum required sensitivity for HT to yield a positive clinical benefit. At specificity values of 100%, 98% and 96%, HT was clinically inferior to standard care for sensitivity values below 82%, 90% and 96% respectively; whilst at a specificity level of 94%, all of the results were clinically inferior to standard care regardless of the level of sensitivity chosen. Based on these results, stage 2 of the analysis focused on exploring sensitivity values between 100% and 82%, and specificity values between 100% and 96%.

**Table 6-1 Incremental effectiveness (total QALY gains and number of new secondary cases across patients part of the *evaluation set* for HT compared to LTHT testing strategy at each diagnostic sensitivity and specificity pair with a rapid test turnaround time**

| HT strategy | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | Diagnostic specificity | | | | | | | |
| | | *100%* | | *98%* | | *96%* | | *94%* | |
| | | *ΔQALY* | *New secondary cases* | *ΔQALY* | *New secondary cases* | *ΔQALY* | *New secondary cases* | *ΔQALY* | *New secondary cases* |
| **Diagnostic sensitivity** | *100%* | **0.469** | **-12** | **0.314** | **-6** | *0.136* | *1* | -0.007 | 7 |
| | *98%* | **0.399** | **-11** | **0.247** | **-5** | *0.076* | *2* | -0.065 | 8 |
| | *96%* | **0.337** | **-10** | **0.183** | **-4** | *0.021* | *3* | -0.125 | 9 |
| | *94%* | **0.269** | **-9** | **0.117** | **-3** | -0.049 | 4 | -0.192 | 9 |
| | *92%* | **0.205** | **-7** | **0.049** | **-1** | -0.108 | 5 | -0.252 | 10 |
| | *90%* | **0.145** | **-6** | *-0.005* | *0* | -0.161 | 6 | -0.306 | 11 |
| | *88%* | **0.078** | **-5** | -0.077 | 1 | -0.236 | 7 | -0.382 | 13 |
| | *86%* | **0.006** | **-4** | -0.140 | 2 | -0.299 | 8 | -0.435 | 13 |
| | *84%* | *-0.063* | *-2* | -0.206 | 3 | -0.359 | 9 | -0.501 | 14 |
| | *82%* | *-0.129* | *-1* | -0.266 | 4 | -0.417 | 10 | -0.564 | 16 |
| | *80%* | -0.191 | 1 | -0.332 | 6 | -0.473 | 11 | -0.625 | 17 |
| | *70%* | -0.519 | 8 | -0.641 | 12 | -0.783 | 17 | -0.921 | 22 |
| | *50%* | -1.194 | 24 | -1.280 | 26 | -1.391 | 29 | -1.518 | 33 |
| *15 min test turnaround time (best case value)* | | | | | | | | | |

Values in **bold** format represent the sensitivity specificity pairs at which HT yields higher QALY gains and leads to a lower number of new secondary infections

Values in *italic* format represent the sensitivity specificity pairs at which HT yields lower QALY gains but leads to a lower number of new secondary infections

The bold black line indicates the threshold at which the diagnostic sensitivity and specificity pairs yield better results on at least one of the clinical outcomes (e.g. higher ΔQALY or lower new secondary cases) [i.e. HT is not clinically inferior]. The cell colours provide an indication of where the results lie on the full spectrum of the outputs observed: solid green is associated with the highest clinical effectiveness outputs (e.g. higher ΔQALY and/or lower new secondary cases); whilst solid red is associated with the lowest clinical effectiveness outputs (e.g. lower ΔQALY and/or higher new secondary cases).

#### 6.3.1.2 Stage 2 – Cost-effectiveness outputs

Table 6-2 reports the incremental cost-effectiveness in terms of INMB results for HT compared with LTHT testing strategy assuming 15-, 90- and 180-minute test turnaround time scenarios. At this stage, the unit cost for HT was assumed to be zero.

Assuming perfect diagnostic accuracy, the INMB for HT compared with LTHT testing strategy was equal to £599,909 when assuming a 15-minute turnaround time; £577,982 at a 90-minute turnaround time; and £535,076 at a 180-minute turnaround time. This indicates that the INMB decreases the longer it takes to receive test results for CDI, while assuming perfect diagnostic sensitivity and specificity.

Once diagnostic specificity was reduced to 96%, the minimum sensitivity for HT to remain cost-effective was 96% at a 15-minute turnaround time, and 98% at both 90- and 180- minute turnaround time. These results indicate that the region of diagnostic sensitivity and specificity where HT is expected to be more clinically and cost-effective reduces as the test turnaround time increases.

Based on the results of stage 1 and 2, for each of the following test turnaround time values, the minimum diagnostic accuracy pairs at which HT is both clinically- and cost-effective were estimated at:

- **15-minute turnaround time** – 96% diagnostic sensitivity and 96% diagnostic specificity;
- **90-minute turnaround time** – 98% diagnostic sensitivity and 96% diagnostic specificity;
- **180-minute turnaround time** – 98% diagnostic sensitivity and 96% diagnostic specificity.

Based on these results, the subsequent stage of analysis focused on summarising the minimum performance specifications and estimating the headroom and threshold unit prices for HT to be cost-effective, sorted by test turnaround time value.

**Table 6-2 INMB for HT compared to LTHT testing strategy at each sensitivity specificity pair assuming zero test cost for HT, sorted by test turnaround time values (15, 90, 180 minutes)**

| HT strategy | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | Diagnostic specificity | | | | | | | | |
| | | 100% | | | 98% | | | 96% | | |
| ***Test turnaround time*** | | *15 min* | *90 min* | *180 min* | *15 min* | *90 min* | *180 min* | *15 min* | *90 min* | *180 min* |
| **Diagnostic sensitivity** | **100%** | **£599,909** | **£577,982** | **£535,076** | **£369,939** | **£319,596** | **£303,845** | *£106,146* | *£90,087* | *£49,603* |
| | **98%** | **£542,445** | **£514,831** | **£485,013** | **£326,192** | **£289,570** | **£271,645** | *£73,285* | *£47,207* | *£9,380* |
| | **96%** | **£505,136** | **£496,729** | **£453,465** | **£286,989** | **£243,111** | **£228,143** | *£48,042* | £3,652 | -£18,645 |
| | **94%** | **£468,668** | **£449,998** | **£410,489** | **£252,306** | **£214,197** | **£183,851** | £3,075 | -£36,668 | -£49,023 |
| | **92%** | **£422,843** | **£392,997** | **£380,708** | **£194,617** | **£169,993** | *£147,554* | -£35,615 | -£68,601 | -£92,634 |
| | **90%** | **£377,578** | **£362,764** | **£339,100** | *£162,391* | £134,103 | £111,918 | -£69,659 | -£103,348 | -£128,604 |
| | **88%** | **£344,772** | **£321,287** | **£312,854** | £120,459 | £89,531 | £83,482 | -£112,850 | -£145,998 | -£165,433 |
| | **86%** | **£293,759** | **£273,130** | *£268,902* | £81,951 | £49,682 | £47,261 | -£146,585 | -£183,775 | -£192,265 |
| | **84%** | **£247,124** | *£235,436* | *£219,646* | £36,798 | £9,792 | £1,202 | -£176,460 | -£211,350 | -£236,039 |
| | **82%** | **£196,151** | *£187,982* | *£170,778* | -£1,929 | -£20,368 | -£36,192 | -£213,339 | -£243,500 | -£269,093 |
| **WTP £20,000 per QALY gained** | | | | | | | | | | |

Values in **bold** format represent the sensitivity specificity pairs at which HT yields higher QALY gains, leads to a lower number of new secondary infections and is more cost-effective

Values in *italic* format represent the sensitivity specificity pairs at which HT yields lower QALY gains, but leads to a lower number of new secondary infections and is more cost-effective

Values in underline format represent the sensitivity specificity pairs at which HT is clinically inferior [i.e. ΔQALY and/or more new secondary cases] but maintains cost-effectiveness

The bold black line indicates the threshold at which the diagnostic sensitivity and specificity pairs yield better clinical outputs on at least one of the clinical outcomes (e.g. higher ΔQALY or lower new secondary cases) [i.e. HT is not clinically inferior] and HT is more cost-effective - sorted by test turnaround time values. The dotted black line indicates the threshold at which the diagnostic sensitivity and specificity pairs yield worse clinical outputs on at least one of the clinical outcomes [i.e. HT is clinically infectior] but maintain cost-effectiveness - sorted by test turnaround time values.The cell colours provide an indication of where the results lie on the full spectrum of the outputs observed: solid green is associated with the highest cost-effectiveness outputs (e.g. higher INMB); whilst solid red is associated with the lowest cost-effectiveness outputs (e.g. lower INMB)

### 6.3.1.3 Stage 3 – Minimum test performance specifications and maximum acceptable prices

| | *Test turnaround time* | | |
|---|---|---|---|
| | *15 minutes* | *90 minutes* | *180 minutes* |
| **Minimum diagnostic sensitivity** | 96% | ↑ 98% | ↑ 98% |
| **Minimum diagnostic specificity** | 96% | ▬ 96% | ▬ 96% |
| **Headroom unit price** [assuming optimal diagnostic accuracy] at WTP £20,000 per QALY gained | £554.36 | ↓ £534.09 | ↓ £494.45 |
| **Threshold unit price** [assuming minimal diagnostic accuracy] at WTP £20,000 per QALY gained | £44.39 | ↓ £43.62 | ↓ £8.67 |

Minimum performance specifications and maximum acceptable unit price at 90 and 180 minutes turnaround time are compared against minimum performance specifications and maximum unit price at 15-minutes turnaround time (baseline value). A grey up arrow is associated with higher minimum performance specification and maximum unit costs compared to results at 15-minutes turnaround time. A yellow bar is associated with minimum performance specifications and maximum unit cost at 90 and 180 turnaround time equal to the respective results at 15-minutes turnaround time. A red down arrow is associated with lower minimum performance specifications and lower maximum costs at 90 and 180 minutes turnaround time.

**Table 6-3 Minimum performance specifications for HT and maximum acceptable prices at each test turnaround time value**

Based on the presented results for stages 1-3, Table 6-3 reports the minimum performance requirements and maximum acceptable prices associated with the HT in order to achieve cost-effectiveness, for the three test turnaround times evaluated. At a 15-minute turnaround time, the minimum diagnostic sensitivity and specificity were both 96%. If we consider the results at 15 minutes as a base case scenario, the lowest minimum sensitivity and specificity which maintained cost-effectiveness increased the longer it took to get test results back. In the context of a test turnaround time of 90 and 180 minutes, the minimum sensitivity for HT to be cost-effective was 98%, whereas the minimum specificity remained 96%. Table 6-3 indicates that a test with slower turnaround time requires a higher sensitivity to maintain a positive clinical benefit and cost-effectiveness compared to standard care, while holding constant a minimum specificity of 96%.

The maximum unit cost for HT to be cost-effective, however, decreased as test turnaround time increased due to a reduction in clinical effectiveness and higher healthcare costs. At a 15-minute test turnaround time, HT can be priced up to £544.36 assuming perfect diagnostic accuracy (i.e. the *headroom price*) and still

be cost-effective at a WTP threshold of £20,000. When diagnostic sensitivity and specificity were reduced to 96% and 96% (the minimum performance requirements at 15-minute test turnaround time identified in stage 2), the maximum price for HT whilst maintaining cost-effectiveness (i.e. the *threshold price*) was approximately £44.39 at a £20,000 WTP threshold.

Assuming a 90-minute turnaround time, the headroom price for a perfectly accurate HT was £534.09 using a WTP threshold of £20,000. If diagnostic sensitivity and specificity were set to their minimum value identified in stage 2 (98% sensitivity, 96% specificity), the maximum price of HT maintaining cost-effectiveness (i.e. the *threshold price*) was £43.62 at a £20,000 WTP threshold. At a 180-minute turnaround time and perfect diagnostic accuracy, the headroom price for HT to be cost-effective was £494.45; once diagnostic accuracy was reduced to its minimally acceptable value (sensitivity 98%, specificity 96%), HT can be priced up to £8.67 and remain cost-effective at a £20,000 WTP threshold.

### 6.3.2 Phase 2: Identification of key model drivers

Under the minimum performance specifications at a 15-minute test turnaround time, HT was clinically more effective compared to the LTHT testing strategy, leading to 0.0205 total incremental QALY gains, despite yielding 3 additional secondary cases of CDI in the general ward[25]. Marginally fewer patients were moved into single rooms comparing HT with minimum performance specifications at 15-minute turnaround time and LTHT testing strategy, whereas more patients were placed in general ward under HT compared to LTHT testing strategy. In this context, 57.95% and 39.70% of patients were placed in single rooms and general ward under the HT testing strategy with minimal performance specifications, respectively, as opposed to 59.39% and 38.70% being placed in single rooms and general ward under the LTHT testing strategy, respectively.

[25] Despite HT leading to more secondary cases compared to LTHT testing strategy, at 15-minutes turnaround time HT yielded QALY gains as faster testing strategies improved short-term patient health (e.g. higher probability of clinical cure, shorter LOS) to a greater extent compared to testing strategies with slower turnaround time. The higher individual QALY gains associated with faster test turnaround time offset the reduction in QALY gains associated with an increased infection spread.

Applying the threshold testing unit cost (i.e. £44.39), HT was less expensive compared to the LTHT testing strategy, leading to an increase in total costs of £410. This, in turn, led to an INMB of £0 per QALY gained at a WTP threshold of £20,000 per QALY gained.

Table 6-4 summarises the incremental clinical- and cost-effectiveness outputs associated with the minimum performance specifications for HT compared to LTHT testing strategy at 15-minutes turnaround time.

This section presents the results of the deterministic univariate sensitivity analysis (see section 0), and scenario analysis (section 6.3.2.2) on the cost-effectiveness outputs associated with the minimum performance specifications for HT at a 15-minute turnaround time compared to LTHT testing strategy. Specifically, these analyses assume an optimal test turnaround time of 15 minutes, the minimum values of diagnostic sensitivity (96%) and specificity (96%), and the threshold price of the test (£44.39) identified in section 6.3.1.3.

**Table 6-4 Capacity constraints-related outputs, new secondary cases of CDI, total and per patient QALY gains, total and per patient costs at baseline for each testing option being evaluated, and incremental differences between HT meeting the minimum performance specifications at 15-minutes turnaround time and LTHT testing strategy**

| | HT – perfect diagnostic accuracy at 15-minutes turnaround time | HT – minimum performance specifications at 15-minutes turnaround time | LTHT testing strategy | Incremental difference (HT – LTHT testing strategy) |
|---|---|---|---|---|
| **Number of occasions a single room was needed and was not available (%)[26]** | 480 (31.46%) | 591 (41.36%) | 586 (40.79%) | 5 (0.57%) |
| **Percentage of patients placed in single rooms[26]** | 67.99% | 57.95% | 59.39% | -1.44% |
| **Percentage of patients placed in cohort bays[26]** | 1.66% | 2.35% | 1.91% | 0.44% |
| **Percentage of patients placed in general ward[26]** | 30.35% | 39.70% | 38.70% | 1% |
| **New secondary cases of CDI** | 28.46 | 43.79 | 40.87 | 3 |
| **Total QALY gains** | 48.07582 | 47.62755 | 47.60704 | 0.0205 |
| **Per patient QALY gains[26]** | 0.08919 | 0.08836 | 0.08832 | 0.00004 |
| **Total costs** | £17,032,005 | £17,574,907 | £17,574,496 | £410 |
| **Per patient costs[26]** | £31,599 | £32,607 | £32,606 | £1 |
| **INMB at £20,000 QALY gained** | | £0 | | |

[26] Across 70 model replications, the size of the cohort entering the model on average 539 patients.

#### 6.3.2.1 Deterministic univariate sensitivity analysis

Figure 6-3 presents the results of the deterministic univariate sensitivity analysis for the most influential model parameters using estimates from the literature and the +/- 25% deviation from baseline values (see Appendix O for full results). Based on these results, the top three most influential drivers of changes in cost-effectiveness for HT compared to LTHT testing strategy were: (1) operating time for multiplex GI panel; (2) LOS for patients positive to CDI receiving test results under slow time-to-diagnosis; and (3) disease prevalence for CDI.

The operating time for multiplex GI panel was the biggest driver of changes in cost-effectiveness, causing the INMB of the HT strategy to vary between £255,620 and -£1,093,096 at the parameter's lower and upper bound (range £1,348,716), respectively. In the context of a longer operating time for multiplex GI panels, both testing strategies yielded more new secondary cases in the general ward and lower QALY gains compared to baseline values. The HT strategy, however, yielded worse incremental clinical outputs (i.e. more new secondary cases, lower QALY gains) compared to the baseline results.

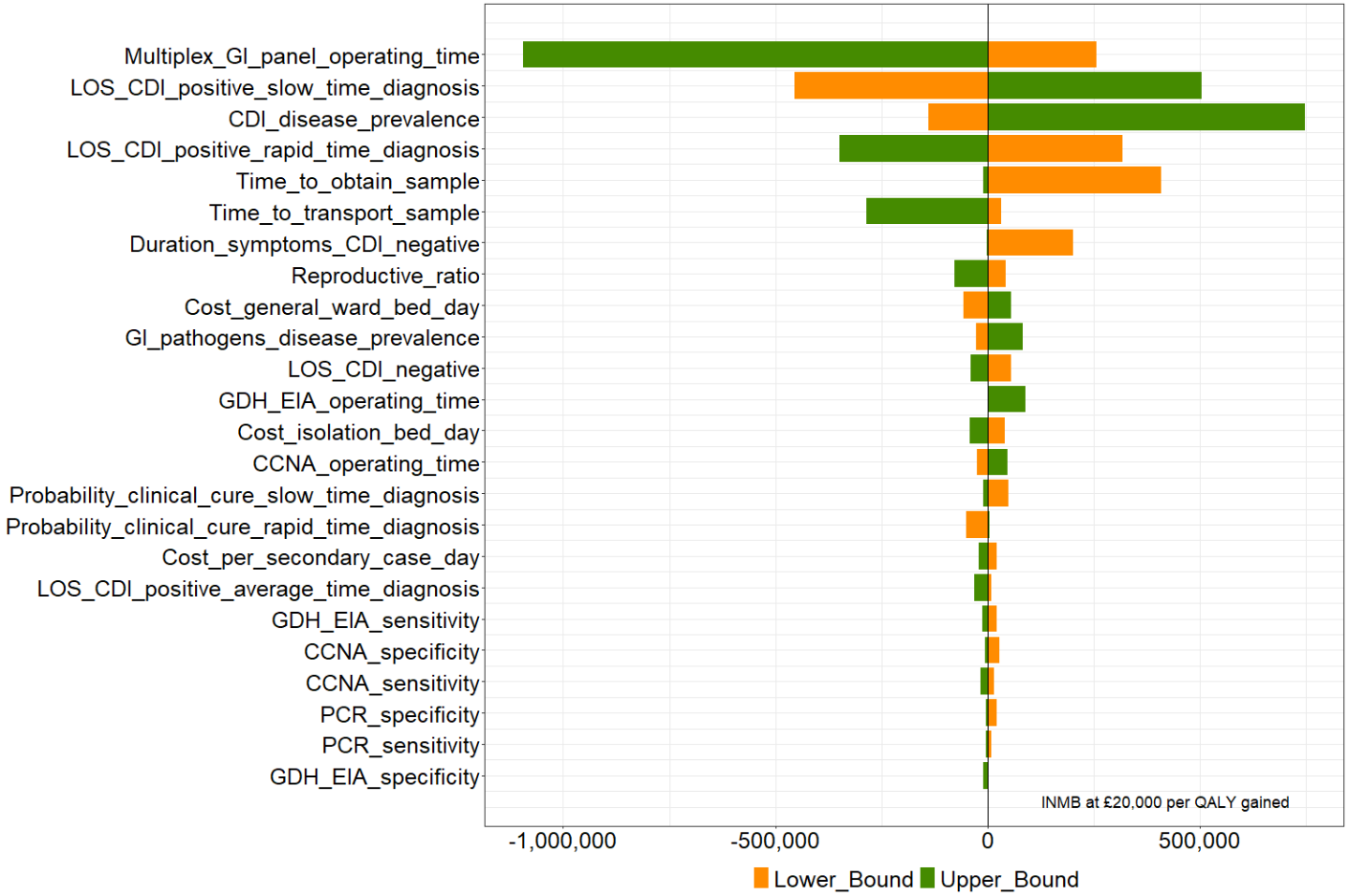
In addition, varying the LOS for patients positive to CDI under slow time-to-diagnosis (i.e. equal or greater to 1.2 days) resulted in a fluctuation of the INMB of £959,857 (-£455,274 and £504,583 INMB at the lower and upper bound, respectively); whilst changes in the LOS for patients confirmed with CDI under rapid time-to-diagnosis (i.e. equal or shorter than 0.5 day) led to a INMB fluctuation of £666,371 (£317,000 and -£349,372 at the lower and upper bound, respectively). In the context of an extended LOS under slow time-to-diagnosis, HT yielded better incremental clinical outputs (i.e. fewer new secondary cases, higher QALY gains) over the LTHT testing strategy compared to the baseline results.

Disease prevalence of CDI was the third most influential parameter causing the INMB to vary by £885,093 (-£139,029 and £746,064 INMB at lower and upper value of disease prevalence of CDI, respectively). In the context of a higher disease prevalence, more confirmed CDI patients would require single room isolation until the end of their hospital stay, thereby increasing the risk of infection spread within the general ward; while more patients truly infected with CDI would experience a worsened quality of life due to CDI.

Based on the results of the deterministic univariate sensitivity analysis using estimates from the literature and +/- 50% deviation from baseline values, the top three parameters having the greatest impact on the cost-effectiveness outputs for HT compared to LTHT testing strategy were:

1. **LOS for patients positive to CDI receiving test results under slow time-to-diagnosis** – causing the INMB of the HT strategy to vary between -£845,650 and £996,568 at the parameter's lower and upper bound (range £1,842,218), respectively.
2. **Disease prevalence for CDI** – resulting in a fluctuation of the INMB of £1,585,739 (-£139,029 and £1,446,710 INMB at the lower and upper bound, respectively).
3. **LOS for patients positive to CDI receiving test results under rapid time-to-diagnosis** – leading to a INMB fluctuation of £1,540,127 (£603,949 and -£936,178 at the lower and upper bound, respectively).

**Figure 6-3 Tornado plot showing the results of the deterministic univariate sensitivity analysis for the most influential model parameters using estimates from the literature and the +/- 25% deviation from baseline values. Results are expressed in terms of INMB at £20,0000 per QALY gained. INMB at baseline is equal to £0 and reflects the cost-effectiveness output associated with the minimum performance specifications for HT at 15-minutes turnaround time**



#### 6.3.2.2 Scenario analysis

Scenario analyses were conducted to assess the impact of structural aspects of the clinical pathway on the clinical and cost-effectiveness outputs associated with the minimum performance benchmarks for HT estimated using the MPS framework (see section 6.3.1.3).

Table 6-5 presents the results of the scenario analysis compared to baseline results. Simulating a typical UK District Hospital (S4) and increasing the availability of multiplex GI panels (S2) were found to have the greatest impact on the cost-effectiveness outputs associated with the minimum performance specifications for HT at 15-minute turnaround time.

The following subsections describe the impact of each structural scenario on clinical- and cost-effectiveness outputs associated with the minimum performance specifications at 15-minute turnaround time, separately.

**Table 6-5 Total new secondary cases of CDI, incremental QALY gains, incremental costs (across cross patients part of the *evaluation set*, INMB between HT meeting the minimum performance specifications at 15-minutes turnaround time and LTHT testing strategy for each structural scenario being explored. Scenarios highlighted in orange have the greatest impact on the cost-effectiveness outputs associated with the minimum performance specifications for HT at baseline.**

| Scenario | New secondary cases | | | QALY gains | | | Total costs | | | INMB at £20,000 per QALY gained |
|---|---|---|---|---|---|---|---|---|---|---|
| | HT strategy | LTHT testing strategy | Δ | HT strategy | LTHT testing strategy | Δ | HT strategy | LTHT testing strategy | Δ | |
| **Baseline results** | **44** | **41** | **3** | **47.63** | **47.61** | **0.02** | **£17,574,907** | **£17,574,496** | **£410** | **£0** |
| **S.1 - Alternative de-escalation protocol for patients confirmed with CDI** | 22 | 20 | 2 | 47.96376 | 47.91891 | 0.04 | £16,965,926 | £16,968,019 | -£2,093 | £2,990 |
| **S.2 - Increased availability of multiplex GI panels** | 37 | 36 | 1 | 47.90054 | 47.80508 | 0.10 | £17,090,958 | £17,239,846 | -£148,887 | £150,796 |
| **S.3 - Reduced availability of single rooms in a Teaching Hospital** | 56 | 53 | 3 | 47.44005 | 47.41238 | 0.03 | £17,843,338 | £17,848,614 | -£5,276 | £5,829 |
| **S.4 - District Hospital** | 14 | 15 | -1 | 48.8649 | 48.73309 | 0.13 | £9,845,706 | £9,996,969 | -£151,263 | £153,899 |
| **S.5 - Second-best parametric distribution for time-to-event patient parameters** | 36 | 34 | 2 | 47.74101 | 47.68916 | 0.05 | £17,973,201 | 18122094.1 | -£148,893 | £149,930 |
| **S.6 - Inclusion of labour costs for running laboratory testing** | 44 | 41 | 3 | 47.63 | 47.61 | 0.02 | 17576194.41 | 17577085.31 | -£891 | £1,301 |

#### 6.3.2.2.1 S.1 – Alternative de-escalation protocol for patients confirmed with CDI

Simulating a more rapid de-escalation protocol resulted in HT yielding a higher INMB (£2,990) compared to the baseline results. Releasing patients confirmed with CDI upon symptoms resolution (rather than at the end of hospital stay) increased the availability of single rooms thereby reducing the number of new secondary cases in the general ward – across both testing strategies. Both testing strategies thus yielded lower incremental QALY lost due to secondary cases and therefore higher incremental QALY gains (0.04 compared to 0.02 at baseline). Under the LTHT testing strategy, however, patients experienced a longer turnaround time to yield final diagnosis which, in turn, led to worse clinical outcomes and thereby lower QALY gains (47.92) versus the HT testing strategy (47.96).

While the isolation costs were lower for both testing strategies under this scenario (£2,094,637 and £2,059,311 for HT and LTHT testing strategy, respectively) compared to baseline results (£2,396,752 and £2,310,641 for HT and LTHT testing strategy, respectively), general ward bed costs were higher. This reflected the longer LOS patients positive to CDI spent in general ward under S1 as opposed to single room isolation, as per the baseline scenario. The HT strategy, however, was associated with lower incremental isolation bed costs versus the LTHT testing strategy compared to baseline results. The marginal reduction in isolation costs therefore lead to higher incremental costs savings (-£2,093) between both strategies compared to the baseline results.

#### 6.3.2.2.2 S.2 – Increased availability of multiplex GI panels

Adding one extra multiplex GI panel to test patients suspected with infectious diarrhoea led to a higher INMB for the HT against LTHT testing strategy (£150,796), compared to the baseline results.

Having a quicker confirmation of infectious diarrhoea due to other GI pathogens expedited the time-to-diagnosis for patients suspected with CDI. This increased the availability of single rooms as patients negative to CDI were more quickly de-isolated compared to baseline scenario. The faster assessment of GI pathogens allowed the full benefit of the HT rapid test to be realised, with HT in this scenario leading to fewer secondary cases within general ward over LTHT testing strategy

(0.51) compared to baseline (3), which resulted in additional cost savings (-£148,887). An additional benefit of shortening the test turnaround time was the increased effectiveness of antibiotic treatment for CDI, which meant that patients positive to CDI also experienced better short-term clinical outcomes (0.10 QALY gain vs. standard care).

#### 6.3.2.2.3 S.3 – Reduced availability of single rooms in Teaching Hospital

Reducing the number of single rooms available in a typical UK Teaching Hospital resulted in HT yielding a higher INMB compared to baseline results (£5,829). Both testing strategies yielded more new secondary cases, lower QALY gains and higher costs in this scenario compared to baseline results. In the context of a reduced availability of single rooms, the HT yielded fewer incremental secondary cases over LTHT testing strategy (2.86) compared to baseline results (3) as patients received a final diagnosis for CDI in a more timely manner as opposed to the comparator testing strategy. As the HT prevented more secondary cases within the general ward, this led to higher incremental QALY gains (0.03) and incremental cost savings (-£5,276) compared to baseline outputs.

#### 6.3.2.2.4 S.4 – UK District Hospital

Simulating a typical UK District Hospital led to a higher INMB for HT over the comparator testing strategy compared to baseline (£153,899). Both testing strategies yielded fewer new secondary cases in the general ward, higher QALY gains and lower costs.

Although a typical UK District Hospital was associated with fewer single rooms where to isolate patients suspected and confirmed with CDI, under S4 a reduced number of patients were tested for CDI thereby there was a reduced demand for single rooms compared to UK Teaching Hospitals. This, in turn, resulted in fewer new secondary cases and thereby higher QALY gains. HT yielded fewer comparative secondary cases for CDI (-1.30) over LTHT testing strategy compared to baseline output (3), and led to higher incremental QALY gains (0.13) and lower incremental costs (-£151,263).

#### 6.3.2.2.5 S.5 – Alternative distributions for time-to-event parameters

Applying the second-best distribution for the time-to-event patient parameters (e.g. LOS, duration of symptoms), resulted in an increase of the cost-effectiveness output for HT compared to LTHT testing strategy (INMB equal to £149,930).

Across both testing strategies, simulating the second-best distribution to time-to-event patient parameters resulted in fewer new secondary cases, higher incremental QALY gains and higher costs. The HT strategy yielded better incremental clinical benefits (i.e. fewer new secondary cases, higher QALY gains) over the comparator testing strategy, compared to baseline results. In this context, HT was associated with a greater reduction in new secondary cases, leading to higher incremental QALY gains compared to LTHT testing strategy.

Applying the second-best distribution had the following downstream implications:

- **lower average LOS for patients confirmed with CDI, and reduced shorter right tail of high LOS values (i.e. fewer extreme high values)** – this, in turn, led to lower isolation costs associated with both testing strategies (£2,228,105 for HT and £2,206,964 for the LTHT testing strategy) compared to the respective baseline values (£2,396,752 and £2,310,640, respectively);
- **higher average LOS for patients negative to CDI on average but shorter right tail** – overall, this led to a higher general ward bed costs associated with both testing strategies (£14,596,509 and £ 14,878,135 for HT and LTHT testing strategy) compared to baseline values (£13,617,912 and £13,617,790 at the base case). The reduction in isolation costs were greater for HT compared to the LTHT testing strategy – the economic benefit of reducing the LOS for patients negative to CDI was offset by a longer turnaround time to receive confirmation of non-infectious diarrhoea under LTHT testing strategy, thereby leading to higher isolation costs for the LTHT testing strategy compared to HT.

#### 6.3.2.2.6 S.6. Inclusion of labour costs for running laboratory testing

Capturing additional labour costs for running laboratory-based tests as part of the standard care strategy results in a marginal increase of the cost-effectiveness outputs for HT compared to LTHT testing strategy (INMB equal to £1,301).

In this context, LTHT was associated with higher testing costs (£52,063 compared to £49,474 at baseline) thereby leading to an increase of cost savings for HT compared to LTHT testing strategy (-£891 compared to £410 at baseline). Although this increase in costs applied only to LTHT testing strategy, the marginal impact of this scenario on the relative cost-effectiveness outputs might be explained by the low labour costs for each testing options as part of the standard care strategy.

### 6.3.3 Phase 3: Sensitivity analyses on the minimum performance specifications

Based on the results of the deterministic univariate sensitivity and scenario analyses (phase 2), additional sensitivity analyses were conducted to assess the impact of the most influential model parameters using estimates from the literature and +/- 25% deviations from baseline values, as well as scenarios on the minimum performance specifications identified in phase 1 (see Table 6-7).

The most influential parameters varied in this phase included: (1) operating time for multiplex GI panel; and (2) LOS for patients confirmed with CDI under slow time-to-diagnosis; whereas the top influential scenarios explored in this phase entailed: (i) simulating a typical UK District Hospital; and (ii) increased availability of multiplex GI panels. Table 6-6 reports the levels of the model parameters varied in phase 3, as well as the main features of the most influential scenarios.

**Table 6-6 Summary of the values of the model parameters varied in phase 3 based on the +/- 25% range, and main features of the most influential scenarios**

| Parameter | Lower bound (+/- 25% deviation from baseline values) | Upper bound (+/- 25% deviation from baseline values) |
|---|---|---|
| Operating time for multiplex GI panel | 30 minutes | 50 minutes |
| LOS CDI positive patients – slow diagnosis | Shape[1]: 0.6491104<br>Scale[1]: 16.60751 | Shape[1]: 1.043657<br>Scale[1]: 38.52885 |
| **Scenario** | **Summary key structural assumptions** | |
| UK District Hospital | • 64 single rooms being available<br>• 850 monthly average of stool samples being tested<br>• Off-site laboratory | |
| Increased availability of multiplex GI panels | • 4 multiplex GI panel testing machines | |

[1] Shape and scale parameters for Weibull distribution

This section presents how the minimum performance specifications at a 15-minute test turnaround time varied depending on the lower and upper values of the most influential model parameters and structural scenarios.

**Operating time for multiplex GI panel**

Applying a higher value of operating time for multiplex GI panels (50 minutes, vs 43 minutes in the base case) resulted in higher minimum performance specifications (99% for both diagnostic sensitivity and specificity), and a corresponding lower threshold unit price (£2) and headroom unit price (£179) compared to baseline values.

Alternatively, simulating a lower value of operating time for multiplex GI panels (30 minutes) led to the same required minimum diagnostic specificity as in the base case analysis (96%), but a lower value of minimum diagnostic sensitivity was accepted (93% vs 96% in the base case). In this context, the additional comparative benefit afforded by a more rapid GI panel (e.g. £255,620 INMB as estimated in phase 2) resulted in a comparatively lower minimum sensitivity required for HT to maintain clinical- and cost-effectiveness. This, in turn, resulted in a higher threshold (£159 vs £44 in the base case), and headroom unit price (£772 vs £554 in the base case).

**Length of stay confirmed CDI patients receiving slow diagnosis**

Simulating an extended LOS for patients confirmed with CDI receiving slow diagnosis resulted in a lower minimum diagnostic specificity (94%, vs 96% in the base case) while maintaining the same required minimum diagnostic sensitivity as in the base case results (96%), and a corresponding higher threshold and (£310, vs £44 in the base case) headroom unit price (£1,031, vs £554 in the base case).

Assuming a reduced LOS for patients confirmed with CDI receiving slow diagnosis led to a higher minimum diagnostic sensitivity (99%, vs 96% in the base case) and diagnostic specificity (99%, vs 96% in the base case). In this context, the reduced clinical and economic benefit of simulating an extended LOS (i.e. -£455,274 INMB as estimated in phase 2) demanded a comparatively higher minimum performance requirements for HT to be clinical and cost-effective. This also resulted in a significantly lower threshold (£10 vs £44 in the base case) and headroom unit price for HT (£144, vs £544 in the base case).

**District Hospital**

In the scenario of a typical UK District Hospital, the minimum performance specifications for HT to maintain cost-effectiveness were lower compared to baseline (minimum diagnostic sensitivity and specificity equal to 92% and 95%, respectively). Following the increased clinical and economic benefit of HT in this context (i.e. £153,899 INMB as estimated in phase 2), a corresponding higher threshold cost was accepted in this context (£84, vs £44 in the base case). Assuming perfect diagnostic accuracy at a 15-minute turnaround time, the headroom price for HT to maintain cost-effectiveness was also found to be higher (£2,818) compared to the baseline value.

**Increased availability of multiplex GI panels**

While simulating an additional multiplex GI panel to test samples suspected of infectious diarrhoea, the minimum performance specifications for HT to maintain clinical and cost-effectiveness were found to be lower (90% and 93% for diagnostic sensitivity and specificity, respectively) than the baseline values. The maximum threshold cost associated with the above minimum performance specifications was also higher than the base case value (£360). In addition, the headroom price for HT was higher (£590) than the baseline value.

**Table 6-7 Minimum performance test requirements and associated maximum acceptable unit cost at 15-minutes turnaround time, and maximum unit cost under perfect diagnostic accuracy, sorted by the low and upper value of the top influential model parameters and structural scenario**

| | Minimum requirements at 15-minutes test turnaround time | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | Baseline parameter value | Multiplex GI panel time-to-result | | Length of stay confirmed CDI patients receiving slow diagnosis | | District Hospital | Increased availability of multiplex GI panels |
| | | Low parameter value ↑ | High parameter value ↓ | Low parameter value ↓ | High parameter value ↑ | ↑ | ↑ |
| Diagnostic sensitivity | **96%** | 93% | 99% | 99% | 96% | 92% | 90% |
| Diagnostic specificity | **96%** | 96% | 99% | 99% | 94% | 95% | 93% |
| Threshold unit price (WTP £20,000) | **£44** | £159 | £2 | £10 | £310 | £84 | £360 |
| Headroom unit price (WTP £20,000) | **£554** | £772 | £179 | £144 | £1,031 | £2,818 | £590 |

The arrows in the headers are associated with the results of phase 2 - an upper arrow indicates that in the context of this parameter value or structural scenario the marginal benefit of HT was increased compared to baseline (i.e. higher INMB). A lower arrow indicates that in the context of this parameter value or structural scenario the marginal benefit of HT was decreased compared to baseline (i.e. lower INMB)

A grey up arrow is associated with higher minimum performance specifications compared to baseline results at 15-minutes test turnaround time. A green up arrow is associated with a higher maximum unit cost for HT compared to baseline value. A yellow bar is associated with minimum performance specifications and maximum unit cost equal to the respective results at 15-minutes test turnaround time. A grey down arrow is associated with lower minimum performance specifications compared to baseline results at 15-minutes test turnaround time. A red down arrow is associated with lower maximum unit cost for HT compared to baseline value.

## 6.4 Discussion

This section discusses the results of the analysis, followed by some methodological considerations on the model analysis and an overview of the model limitations.

### 6.4.1 Identified minimum performance specifications

Based on the presented MPS framework, a minimum diagnostic sensitivity and specificity of 96% is required for a hypothetical POCT for CDI to be cost- effective compared to standard care. Provided that a new test achieves these minimum levels, and could yield results in 15 minutes, a reasonable maximum unit price for that test would be approximately £44. Assuming that perfect diagnostic accuracy were technically achievable, manufacturers could charge a maximum unit price of £554 for the new test to be cost-effective. In the context of a test turnaround time of 180 minutes, a higher minimum diagnostic sensitivity is required to maintain clinical and economic benefit (98%, vs 96% at a 15-minute turnaround time), whereas the minimum diagnostic specificity required remains at 96%. The associated maximum unit price for a test meeting these performance specifications would be £9.

The robustness of the minimum performance specifications identified using the MPS framework at a 15-minute turnaround time (i.e. with all model parameters at baseline) was tested against the top drivers of cost-effectiveness changes. Depending on the value of the top influential parameter and/or structural scenario being investigated, the minimum diagnostic sensitivity for HT to maintain clinical and cost-effectiveness varied between 90% and 99%, whereas the minimum diagnostic specificity ranged between 93% and 99%. The corresponding threshold cost ranged between £2 and £360, whereas the headroom cost (for a perfectly accurate test) varied between £144 and £2,818.

In this example, diagnostic specificity was a greater driver of clinical utility and cost-effectiveness compared to diagnostic sensitivity – a reduced diagnostic specificity required a higher minimum diagnostic sensitivity for HT to maintain a positive clinical benefit (see 6.3.1.1). Based on the low prevalence of CDI in this case, incorrectly diagnosing a proportion of healthy patients as CDI positive had a greater impact than missing a few truly positive CDI patients. In addition, the test turnaround time had a substantial impact on the expected clinical and

economic benefit a POCT could afford – the longer the time to receive results for CDI, the lower the maximum unit price for the hypothetical rapid test to maintain cost-effectiveness, while more rigorous minimum performance benchmarks were required to yield a clinical and economic benefit. Test manufacturers should therefore balance any delays in test turnaround time with improvements in analytical and diagnostic accuracy in order for the test to provide sufficient clinical benefits to patients and the hospital workflow. The substantial knock-on impact of the test turnaround time on clinical and economic outcomes can be explained by the fact that patient short-term clinical outcomes (e.g. LOS, treatment effectiveness) were dependent on the time point within which patients receive final diagnosis – as discussed in Chapter 5.

Minimum performance requirements for HT to be more clinically and economically effective than standard care were high. This is largely due to the high performance of the standard care comparator in this case – running the two-stage testing algorithm is a cost-effective option for diagnosing CDI (200, 229), as a highly sensitive GDH screening test rules out negative patients with results being quickly communicated to the clinical team (199); only positive patients then fully undergo the testing pathway. While the identified minimum performance benchmarks could be considered difficult to attain for test manufacturers, recent developments within the innovation pipeline suggest that new rapid diagnostics for CDI could partially meet the minimum performance requirements identified using the MPS framework in certain parameters and scenarios explored in phase 3 (i.e. when the multiplex GI panels can yield results faster, or the number of available panels can be increased, or in the context of UK District Hospitals). Recently, new rapid ultrasensitive toxin tests have shown promising results in the detection of *C. difficile* toxin A/B as standalone tests (308, 309). Three ultrasensitive toxin tests are currently undergoing assay development processes and regulatory approval, including :(i) the *Singulex Clarity C.diff toxin A/B assay* (Singulex Inc., Alameda, CA, USA); (ii) the *single-molecule array (Simoa)* technology (Quanterix, Inc., Lexington, MA, USA); and (iii) *MultiPath* (First Light Diagnostics, Chelmsford, MA, USA) (308). Among the new ultrasensitive tests being developed, the Singulex Clarity C.diff toxin A/B assay had an estimated diagnostic sensitivity of 96.3% and a diagnostic specificity of 93% when compared to CCNA (309).

Assuming that a new test were able to meet the minimum performance benchmarks, further considerations should apply in relation to the maximum unit price manufacturers can charge for a new POCT. Unit prices for POCTs are usually more expensive than laboratory testing due to higher costs of consumables, and the impossibility of achieving economy of scales as opposed to laboratory testing (310). Assuming a total of 8208[27] patients being tested for CDI in secondary care across the UK every year, were the developers able to secure a portion of the market equal to 20% and 40%, they could be expected to earn £722,304 and £1,444,608 over the course of 10 years, respectively – assuming the company would price the new test equal to £44 (i.e. threshold cost at 15-minutes turnaround time). Were the cost of developing and producing a new diagnostic test for CDI lower than the expected return for investment, then developing new tests for CDI would offer sufficient investment for manufacturers. For further details on the calculations and sources needed to calculate the expected return of investment see Appendix O.

Were a manufacturer to charge up to £100 as a unit test price, which provides results in 15 minutes, minimum performance requirements for the test to maintain clinical and cost-effectiveness would likely be higher. Greater clinical benefits associated with a test are therefore required to ensure that a more expensive test still represents a cost-effective strategy. In the context of an extended turnaround time (i.e. 180 minutes), however, the maximum unit price for a test to maintain clinical and cost-effectiveness (£9) is unlikely to provide test developers with sufficient return for investment, thereby reducing the commercial viability of developing new POCTs.

Based on the results of sensitivity analyses conducted in phase 2 using estimates from the literature and +/- 25% deviation from baseline values, the operating time for multiplex GI panels and LOS for patients confirmed with CDI receiving slow diagnosis had the biggest impact on the relative cost-effectiveness of a hypothetical rapid test for CDI over standard care testing strategy. In addition, simulating a District Hospital (S.4) and assuming an additional multiplex GI panel as part of the routine testing workflow (S.2) were found as the most influential scenarios to consider while defining minimum desirable performance

[27] Total number of patients tested for CDI in secondary care is based on the COMBACTE-CDI survey data taking into account both UK District and Teaching Hospital (see Chapter 5).

specifications based on cost-effectiveness considerations. The following paragraphs explain the mechanisms by which the most influential parameters and scenarios affect the minimum performance specifications for HT.

In clinical settings where patients suspected with infectious diarrhoea are tested simultaneously for CDI and other GI pathogens, if faster time-to-diagnosis cannot be maintained (e.g. patients have to wait longer for results on multiplex GI panels due to slower machines and reduced availability of panels), HT strategy leads to fewer single rooms being available and worse clinical outcomes (i.e. more secondary cases and lower QALY gains) thereby leading to a reduced INMB and more rigorous performance requirements compared to baseline results. Under the HT testing strategy, one test result is needed to confirm the diagnosis of CDI, whereas the LTHT testing strategy requires several test results to rule in the diagnosis of CDI. In the standard care arm, having to wait longer for 'other GI pathogens' diagnosis has a marginal impact since the standard care process to yield test results for CDI takes longer compared to HT testing strategy. On the contrary, in the HT testing arm, diagnosis for CDI can be achieved more quickly and it is associated with short-term clinical and economic benefits for the HT strategy. The extended time to wait for test results for other GI pathogens in the HT testing arm offsets any clinical and economic benefits associated with the HT rapid test for CDI if a quick final diagnosis cannot be achieved. The full incremental benefit of the HT strategy is therefore achieved only if the rapid time-to-diagnosis can be maintained.

The relative importance of the operating time for multiplex GI panels and the number of available testing machines suggests that developing multiplex testing panels with the ability to detect several GI pathogens responsible for infectious diarrhoea could lead to a substantial clinical and economic benefit. Detecting simultaneously *C. difficile* toxins and other GI pathogens might lead to faster diagnosis with the potential of de-isolating non-infective patients quickly. If future tests could be developed to detect both *C. difficile* toxins and other GI pathogens, the MPS framework presented in this chapter could be applied to define the minimum performance specifications for a *C. difficile*-specific POCT embedded within a multiplex GI panel strategy.

In addition, the model results were also highly sensitive to changes to the LOS for patients positive to CDI receiving slow diagnosis. Across both testing

strategies, simulating an extended LOS increased the risk of infection spread within the general ward due to a reduced availability of free single rooms, whilst it increased the individual QALY loss as the CDI utility was assigned for longer (see Chapter 5 for detail). Any variation in the LOS associated with slow time to diagnosis, however, solely affected testing strategies with extended turnaround time (i.e. LTHT testing strategy in this context) – as patients receiving rapid confirmation of CDI were assumed to experience LOS under rapid time-to-diagnosis instead. In the context of an extended LOS under slow time-to-diagnosis, HT therefore yielded better incremental clinical outputs (i.e. fewer new secondary cases, higher QALY gains) over the LTHT testing strategy compared to the baseline results. This, in turn, led to lower minimum performance specifications for HT to maintain cost-effectiveness compared to LTHT testing strategy.

Simulating a typical UK District Hospital with fewer isolation rooms and reduced number of patients being tested for infectious diarrhoea was found as one of the most influential scenarios on cost-effectiveness changes. A key driver of these results was the fact that, under the District Hospital scenario, samples were transported to an off-site laboratory thereby leading to an extended time-to-diagnosis for LTHT testing strategy only, whereas testing under HT strategy occurs at point-of-care (as per baseline scenario). Patients tested with HT strategy thereby received confirmation of the diagnosis of CDI more rapidly, leading to the higher QALY gains and cost savings compared to the LTHT testing strategy. Since HT was positioned at point-of-care and did not require transportation to the laboratory, the clinical and economic benefits associated with HT rapid diagnosis could be maintained in this scenario. Failure to consider the time for transporting samples to the laboratory could offset any clinical and economic benefits associated with a new and accurate diagnostic test for CDI – despite the test itself being rapid.

### 6.4.2 Pragmatic approach to define minimum performance specifications – key considerations

Based on the survey results presented in Chapter 4, a need for more accurate and rapid diagnostic tests for CDI emerged. The survey, however, did not attempt to elicit respondents' views on the minimum and desirable diagnostic accuracy for new tests to address this unmet clinical need. To this end, using established methods of cost-effectiveness analysis and early economic modelling (discussed in Chapter 1 and Chapter 2), a novel three-stage framework was developed to identify the minimum test specifications for three performance dimensions (diagnostic sensitivity, diagnostic specificity and test turnaround time), and the associated maximum test costs, based on clinical- and cost-effectiveness model outputs. Stage 1 focused on identifying the region of diagnostic accuracy that resulted in HT being clinically superior versus standard care (i.e. fewer new secondary cases within the general ward or higher incremental QALY gains) at a 15-minute turnaround time, whilst stage 2 aimed to rule out performance levels where HT was not expected to be cost-effective, despite the test itself being costed at zero, at a range of pre-specified test turnaround times. Based on the results from the previous stages, stage 3 identified the minimum diagnostic accuracy which maintained a positive clinical benefit (stage 1) and cost-effectiveness (stage 2), as well as the associated maximum unit prices for the HT to be cost-effective (i.e. headroom and threshold price), for each pre-specified test turnaround time value.

The presented MPS framework represents the first application of early economic evaluation methods as a means of deriving the minimum performance specifications as set out in TPPs. A recent NICE-commissioned early economic model estimated the costs and health benefits of having a hypothetical COVID-19 POCT which met pre-defined minimum and desirable performance specifications for diagnostic accuracy and time-to-result as presented in a TPP issued by the UK MHRA (164). This model, however, did not aim to directly inform the TPP specifications, but rather to estimate if a hypothetical test meeting pre-established minimum and desirable performance benchmarks would be cost-effective (164). It is unclear, in fact, how the requirements listed in the MHRA TPPs were derived. In line with general reporting limitations identified in the methodology review conducted in Chapter 2, the MHRA TPPs for a COVID-19

POCTs similarly failed to adequately report the sources or justification for the presented performance specifications. The analysis presented in this chapter, meanwhile, aimed to derive *de novo* minimum performance specifications using the NICE WTP threshold applied to modelled cost-effectiveness outputs, via a series of interlinked two- and three-way sensitivity analyses. In the context of further stakeholder consultations, these results will provide the foundation for minimum and optimal performance specifications for rapid tests for CDI within a future TPP.

The minimum performance specifications estimated using the MPS framework (phase 1) assumed that the model parameters were held at their baseline values. Due to the (inevitable) uncertainty surrounding the minimum performance requirements estimated with the MPS framework, it is recommended to test the robustness of the minimum performance specifications before presenting the modelling results to a stakeholder audience in the context of TPP development. To this end, extensive deterministic univariate sensitivity and scenario analyses were conducted to test the impact of different model parameters and scenarios on the cost-effectiveness outputs. This helped to: (i) pragmatically identify the main drivers of cost-effectiveness changes (as discussed in the previous section); (ii) narratively draw inferences on the impact of each parameter and structural scenario on the minimum performance requirements; and (iii) suggest key areas for future research.

In the context of parameter values and scenarios where HT led to higher economic and clinical benefits compared to baseline results, lower minimum performance specifications and a higher maximum unit price were required for HT to maintain cost-effectiveness. For example, in the context of a higher disease prevalence for CDI using +/- 25% deviation from baseline value, the benefit of having a rapid POCT for CDI over a multi-stage testing algorithm was found to be greater than at baseline (£746,064 vs. 0[28] INMB). This means that, in countries where the prevalence of CDI is higher than in the UK, developers could charge more for a new POCT for CDI with marginally less stringent minimum

[28] Under the minimum performance specifications at 15-minute test turnaround time, the increased benefits associated with HT were offset by the higher costs since the testing unit cost is set equal to the maximum cost at which HT remains cost-effective (i.e. *threshold cost*). This, in turn, led to an INMB of £0 at a WTP threshold of £20,000 per QALY gained.

performance benchmarks than the baseline performance specifications identified using the MPS framework in phase 1. This finding might be particularly of interest to manufacturers in Eastern European countries, as the COMBACTE-CDI study has found a higher prevalence of CDI compared to Western European countries (253). Conversely, in the context of parameter values and scenarios where HT yielded a lower INMB compared to the baseline analysis, the minimum performance requirements for HT to maintain cost-effectiveness were higher, while a lower threshold unit price is tolerated for the test to remain cost-effective. For instance, in the case of delays in obtaining stool samples for patients, the expected clinical and economic benefits of having a rapid POCT are reduced compared to baseline as it takes longer to de-isolate patients negative to CDI due to a delayed confirmation of non-infectious diarrhoea. This results in higher minimum performance requirements for HT to remain cost-effective, while a lower threshold unit cost is required to maintain cost-effectiveness.

An additional benefit of conducting extensive sensitivity analyses lies in identifying the areas of greatest uncertainty expected to drive cost-effectiveness results – these findings can then help inform future research priorities. In this example, the model results were highly sensitive to changes in the LOS for patients positive for CDI. This reflects a high uncertainty surrounding the time-to-event parameters for patients positive to CDI, as well as their key relevance in driving the clinical and cost-effectiveness outputs. As the time-to-event parameters in this study were based on one study with 126 patients (311), more evidence is needed to confirm the impact that early treatment, via early diagnosis, has on patient LOS and additional short-term clinical outcomes. Once new evidence on these key parameters is available, the presented decision model could be updated and re-run to derive up-to-date performance specifications (2). This chimes with the concept of TPPs being 'living documents', which should be updated once new evidence becomes available.

Results of a deterministic univariate sensitivity analysis, however, might vary depending on the arbitrary range of parameter values explored (312). To overcome this limitation, the selection of the range for univariate sensitivity analysis should be based on consensus with stakeholders in the context of TPP development and be clearly motivated. Alternatively, novel approaches for conducting univariate sensitivity analyses have been recently suggested – such

as the probabilistic one-way sensitivity analysis (POSA) (312). This method captures correlations between model parameters and the probability of a given parameter to take a specific value using a two-level Monte Carlo approach (312).

### 6.4.3 Model strengths and limitations

The model presents a significant methodological improvement from the published decision models for CDI diagnostics (see rapid literature review in Chapter 5). As opposed to published decision models in this context, this model has the functionality to:

- **capture the sequence of events each individual patient experiences within the clinical pathway** – the model structure features the key events and individual trajectory each patient suspected and confirmed with CDI experiences (e.g. suspicion of CDI, presumptive isolation, de-escalation of infection control measures, treatment administration) within the clinical pathway based on their individual characteristics (e.g. symptoms duration). In addition to this, the model accounts for important recurring events patients confirmed with CDI experience over time, such as clinicians' assessing resolution of symptoms and administering additional antibiotic treatment in case patients remain symptomatic.
- **simulate the processes samples undergo as part of the testing pathway** – this is the first model to account for the key processes happening to patients and their stool samples being tested simultaneously within the testing pathway (e.g. transporting the sample to the laboratory, sample preparation, reviewing test results) in such granular detail. This means that the model is able to: (i) capture the structural aspects underpinning the testing pathway which might influence the time-to-diagnosis for CDI (e.g. off-site laboratory, number of samples being tested per machine run); and (ii) assess the impact of receiving a faster or slower diagnosis of CDI on patients' health, treatment effectiveness and availability of single rooms – while accounting for such structural aspects.
- **account for capacity constraints patients face within the clinical pathway** – an additional novel aspect of this model lies in explicitly accounting for the limited availability of single rooms in which isolate patients suspected and confirmed with CDI, as well as time constraints for testing samples in the laboratory. The model structure therefore resembles

as much as possible the conditions and waiting times patients suspected with CDI face once they enter a busy hospital. In addition to this, the model also explored the downstream consequences that a new rapid test could lead to on the infection control infrastructure and, consequently, on the infection spread within the general ward.

In addition to the above, the structure of the model is sufficiently flexible to explore the impact of alternative clinical assumptions and scenarios. Although the model structure reflects the main aspects of the clinical pathway in place at LTHT, the model was programmed in a way that alternative clinical practices could be easily simulated (e.g. alternative de-escalation protocol, reduced availability of single rooms). This model therefore could be of use for future economic evaluations of diagnostic tests for CDI across care settings with different clinical practices and availability of single rooms. The model coding and technical documentation is available upon request, allowing future modellers to use and build on this model in order to explore different research questions, or to update the data sources once new evidence becomes available.

The presented findings, however, should be considered in line with the model limitations, which fall into the following four categories: (i) clinical assumptions; (ii) model parameterisation (iii) model analysis; and (iv) model validation.

**Clinical assumptions**

The model assumes that there is no risk of co-infection between CDI and other GI pathogens responsible for infectious diarrhoea (CA 2). Based on a multicentre evaluation of the *BioFire FilmArray* ® GI panel, an additional GI pathogen was found in 31.5% of the samples positive to CDI (303). Recent USA-based prospective studies, however, have found no significant differences in terms of CDI severity, treatment effectiveness, recurrences and length of hospital stay between patients co-infected and patients confirmed with CDI only (313, 314). It is unclear therefore if simulating the risk of co-infection would have changed the clinical management and short-term outcomes for patients confirmed with both CDI and other GI pathogens. This simplifying assumption was therefore adopted as there is currently a lack of agreement as to how to treat patients who tested positive to more than GI pathogens. In addition to this, the focus of this model lied in the diagnosis of patients positive to CDI only, rather than every GI pathogen responsible for infectious diarrhoea. Were the risk of co-infection to be

captured within the model, this is not expected to have a significant impact on the incremental number of new secondary cases or QALYs gains between HT and LTHT testing strategies. Simulating patients confirmed with both CDI and other GI pathogens would not reduce the availability of single rooms as patients confirmed with infectious diarrhoea are assumed to remain in single room isolation – regardless of the pathogen being detected (CA 3). Although patients positive to other GI pathogens are released more rapidly from isolation compared to patients confirmed with CDI (e.g. 2 days as opposed to the remaining LOS, respectively), this change would affect both HT and LTHT testing strategy equally – thereby the impact on the incremental clinical and economic benefits, and consequently on the minimum performance requirements for HT, is expected to be marginal.

Due to the paucity of data on the impact of early and better diagnosis on patient long-term outcomes (e.g. survival, quality of life), a core clinical assumption underpinning the model structure is that patients truly positive to CDI (TP and FN cases) do not experience a risk of disease recurrence (CA 10 and 11). Based on discussions with clinical experts, it was decided against modelling the impact of early and more accurate diagnosis on the probability of disease recurrence due to paucity of data in this context, as well as the complexities underpinning the treatment decisions for patients with recurrent CDI. Patients with recurrent CDI, however, are typically associated with higher healthcare costs, longer LOS and worse health-related quality of life compared to patients with a first episode of CDI, according to a UK-based multicentre non-interventional study (195). As 20-30% of patients confirmed with CDI experience a disease recurrence (182-184), this generous clinical assumption might have underestimated the economic and clinical burden associated with CDI. Were recurrence events to be included, HT would be expected to yield higher incremental QALYs compared to LTHT testing strategies as rapid testing strategies are typically associated with an overall decrease in recurrence rate over a month based on the prospective time-series study conducted in France (n=126) (260) (see Chapter 5 for detail). In this context, simulating risk of recurrence for patients confirmed with CDI is expected to favour HT thereby leading to lower performance requirements compared to baseline results. Further studies, however, are required to confirm the positive impact of early diagnosis in reducing the risk of disease recurrence for CDI

patients before safely introducing this clinical assumption within the decision model.

An additional limitation lies in assuming that, in the context of single rooms and cohort bays at full capacity, confirmed cases remain in the general ward until hospital discharge (CA 4). In the model, the simplifying assumption was adopted that patients confirmed with CDI could not be released from single room isolation where other patients with a higher priority needed to be isolated. In clinical practice, however, decisions around when to de-escalate infection control measures are based on a continuous assessment of the symptoms duration of patients being isolated, the risk of infection spread and up-to-date availability of free single rooms – for example, if no single rooms are available, lower priority patients (i.e. asymptomatic patients confirmed with CDI) would be released from single room isolation to accommodate for higher-priority patients. Simulating the continuous bed allocation decisions, and every possible factor having an impact on them, however, would have significantly increased the complexity of model structure, and therefore extended the model running time. The time-to-next event progression and individual feature of DES modelling technique make the task of continuously moving patients confirmed with CDI between general ward and single rooms based on the current availability of free single rooms computationally intensive.

In addition, the model assumed that patients confirmed with other GI pathogens would require single room isolation (CA 3). In clinical practice, however, this decision is dependent on the pathogen being detected – for example, a patient positive to *Campylobacter* does not require single room isolation. This simplifying assumption was adopted since the focus of the model was on the diagnosis of CDI, rather than the other GI pathogens responsible for infectious diarrhoea. Simulating how the escalation of infection-control measures would have changed as a result of detecting each specific GI pathogens would have increased the model complexity and computational burden – without providing stakeholders in the context of TPP with valuable information. The result of this simplifying assumption is that the model may have underestimated the availability of free single rooms thereby leading to an increased risk of infection spread within the general ward. If the complexities of specific GI pathogens were to be captured, in the context of a reduced availability of single rooms, HT would have led to higher

incremental QALY gains compared to LTHT testing strategies as faster testing strategies are associated with fewer new secondary cases and, consequently, lower minimum performance requirements compared to baseline results.

In addition, this model did not capture the follow-up laboratory and imaging testing patients positive for CDI need to undergo – as per common practice among the published decision models for CDI diagnostics. It is expected that, were follow-on tests included, this would favour the LTHT testing strategy – as HT with reduced diagnostic accuracy would lead to more FP cases compared to LTHT testing strategy, and therefore more patients may have needed additional follow-up laboratory and imagining testing. As soon as the diagnostic specificity of HT drops below that of LTHT testing strategy, having more testing costs related to FP cases would result in the HT having higher costs than standard care. The minimum performance specifications for HT would therefore be expected to be slightly higher compared to baseline results to maintain cost-effectiveness.

An additional limitation lies in assuming that clinicians fully adhere to test results when deciding when to start or stop administering antibiotic treatment (CA 5). The model therefore failed to capture the downstream implications of clinicians misinterpreting or overriding test results. This limitation, however, is common among early economic models for tests where full compliance to test results and perfect implementation of the new tests is often assumed (29). Were the impact of interpretation error to be simulated in the model, incorrectly misinterpreting a negative test result for CDI as positive would be expected to have a greater knock-on impact on infection spread than misinterpreting a positive test result CDI as negative. This might be explained by the low disease prevalence for CDI, as well as the reduced availability of free single rooms to isolate suspected and confirmed patients due to the higher FP cases. Difficulties in interpreting test results and taking the correct course of action are particularly relevant in the context of CDI as final diagnosis is usually confirmed by the combination of multiple test results. Future iterations of this model should therefore attempt to capture the downstream consequences of misinterpreting test results for CDI on patients' health, infection spread and costs.

**Model parameterisation**

In the model, the monthly demand for stool testing (i.e. the number of patients entering the model) is assumed to be independent of seasonal influences and

other external factors. Retrospective cohort studies conducted in the USA, however, found seasonal variability in the incidence of hospital-acquired CDI associated with an increased incidence of pneumonia and influenza during the winter months (315, 316). In addition, to estimate the monthly demand of stool testing, an exponential distribution was applied – where a constant rate of new patient arrivals is assumed. These simplifying assumptions might have underestimated the number of patients entering the model leading to an increased availability of free single rooms for patients suspected with CDI. In case of a surge of patients with suspected infectious diarrhoea, more patients would be tested and there would be fewer single isolation rooms available. In this context, the ability to diagnose CDI rapidly using the HT testing strategy would be associated with greater economic and clinical benefits compared to standard care. In this scenario, HT would be associated with lower minimum performance requirements to remain cost-effective compared to baseline values. Whilst in periods of lower incidence of infectious diarrhoea, there would be a reduced demand for single rooms thereby the associated benefits of receiving a fast diagnosis for CDI using HT would be expected to be minimised as standard care testing strategy would lead to fewer new secondary cases. In this scenario, the minimum performance requirements associated with HT would be expected to be higher compared to baseline values.

Limited data informed the hospital configuration (e.g. number of single rooms available), which was based on the COMBACTE-CDI survey. Although survey results were discussed and validated with clinical experts, it is unclear if the survey findings are reflective of the NHS hospitals. In addition, estimates of the short-term impact of early diagnosis on patients confirmed with CDI are based solely on single-center prospective time-series study (n=126) (260) – further larger, randomised and multi-center studies are thereby required to confirm the findings of this study.

While comparing CCNA against TC, the diagnostic accuracy for CCNA is likely to be underestimated as CCNA detects toxins, while TC assesses the presence of the toxigenic organism (308). In addition, independence between sequential tests (e.g. GDH EIA, PCR and CCNA) is assumed due to the paucity of data on the diagnostic accuracy of each test run as part of the LTHT testing algorithm. Having more tests run as part of testing algorithm, however, increases the overall

diagnostic accuracy (213) – thus the overall diagnostic accuracy of the comparator testing strategy might have been underestimated. If this turned out to be the case, then the model may currently over-estimate the incremental benefits of HT vs. standard care, and thus underestimate the minimum performance specifications for HT. Future evaluation studies are required to determine the true diagnostic accuracy of CCNA by using a relevant reference standard, and to evaluate the diagnostic accuracy of multi-step testing algorithms for CDI. Future iterations of this model can then accommodate for such new evidence once it emerges.

**Model analysis**

Small incremental QALY gains between the testing strategies under evaluation were observed. This is common among economic evaluations for diagnostic tests given the indirect impact tests have on patient health, which is usually mediated through treatment effectiveness (21). In addition, since the long-term consequences of treatment on patient survival and quality of life were not captured in this model due to paucity of data, the QALY impacts would be expected to be smaller, and subject to significant uncertainty. The paucity of data on the test-treatment pathway raises questions as to whether QALYs are the appropriate metric to sufficiently capture the impact diagnostic tests might have on the care pathway as well as the infection-control infrastructure. This is an issue common to economic evaluations of healthcare interventions aiming to reduce antimicrobial resistance (317), or infection spread (318). A potential way forward, in cases where an alternative clinical outcome is of primary importance (e.g. number of new secondary infections averted, reduction in inappropriate antibiotic prescriptions), could be conducting a WTP elicitation exercise with clinical stakeholders and, subsequently, using the elicited WTP to derive minimum performance specifications.

An additional challenge lies in accounting for the individual and intergenerational health impact of inappropriate antibiotic treatments on antimicrobial resistance (318). Due to limited data, the model did not directly capture the increased risk of becoming colonised with *C. difficile* organism upon receipt of inappropriate antibiotic treatment for CDI (206, 319). Instead, an arbitrary utility decrement was applied in such cases to capture the downstream consequences of receiving incorrect antibiotic treatment. Although this parameter had a marginal impact on

model results based on the findings of the deterministic sensitivity analysis conducted in phase 2, the model would benefit from better data on the downstream impacts of inappropriate antibiotic treatment.

In the model, an indirect approach was taken to capture the impact of possible infection spread, based on calculating the number of new secondary CDI cases. This approach, however, ignores the possible infection spread within cohort bays, as well as the risk of infection spread due to other GI pathogens. Whilst there is a risk of infection spread within cohort bays (for those patients who do not have CDI), there is currently a paucity of high quality data on the impact of grouping patients on the infection transmission (320). In this context, HT would be expected to lead to an increased availability of single rooms compared to standard care due to faster time-to-diagnosis which, in turn, would offset the increased infection rates were the abovementioned additional sources of transmission be captured within the model. This, in turn, would potentially lead to fewer secondary cases, higher QALY gains associated with HT, and therefore lower minimum performance requirements for HT and higher threshold costs compared to baseline values.

Finally, sampling uncertainty was not captured within the model as a PSA was not conducted. While running a PSA is good modelling practice (280), conducting a PSA for each diagnostic accuracy pair and test turnaround time values explored as part of the MPS framework and the subsequent phases of analysis would have significantly increased the model computational burden and running time. This pragmatic consideration was particularly relevant in the context of the complex resource-constrained DES model presented in Chapter 5. Instead, it was decided to focus on how to translate the modelling results into TPP performance specifications, while also identifying what factors within the pathway might affect the cost-effectiveness of a hypothetical rapid test. In addition, the biggest drivers of changes in the cost-effectiveness of HT against LTHT testing strategy were the performance specifications for the hypothetical rapid test (e.g. diagnostic accuracy, turnaround time and test price). Since the purpose of this evaluation was to estimate minimum requirements for these key parameters, for which no data is currently available, it would be difficult to apply any meaningful distributions around such unknown parameters. A published early economic model on hypothetical new typhoid fever diagnostics conducted a PSA while

holding constant the minimum diagnostic accuracy and maximum price derived with threshold and headroom analyses (77). This, however, begs the question as to whether sampling uncertainty was fully captured if the biggest drivers of cost-effectiveness changes are held constant.

Compared to past early economic models, the model presented in Chapter 5 attempted to minimise the impact of *first-order* uncertainty on the outcomes of interest by running the model 70 times – with each run using a different random number sequence. In addition, extensive deterministic sensitivity and scenario analysis were conducted to explore the impact of varying either one or two parameters, or certain key structural scenarios (e.g. applying different distributions on time-to-event parameters). Nevertheless, further iterations of the model could include a PSA while running the MPS framework and additional sensitivity analyses.

**Model validation**

Given the early and exploratory nature of this decision model, it was not possible to externally validate the modelled outputs. This is a common challenge hindering the applicability of early economic models for tests (29), especially in the context of treatment effectiveness. In addition, due to a paucity of data on the in-ward infection spread of CDI within LTHT, it was not possible to externally validate the number of secondary cases for CDI associated with LTHT testing standard care. Although UK data is available on the infection spread of CDI, given the peculiar nature of the testing strategy in place at LTHT compared to the more common testing options across the UK (i.e. testing with EIA for GDH and toxin A/B – see Chapter 4), the limited data from LTHT might hinder the external validation of the clinical outputs. None of the published decision models for CDI diagnostics which captured the infection element of CDI externally validated their clinical outputs with published data. Nevertheless, all possible efforts were undertaken to ensure the internal and face validity of the model (see Chapter 5). Attempts at external validation should be conducted as soon as relevant evidence emerges, to increase the model applicability.

## 6.5 Chapter summary

- In this chapter, the model analysis was conducted to identify the minimum performance specifications (e.g. diagnostic sensitivity and diagnostic specificity, test turnaround time) and associated maximum acceptable unit price for hypothetical rapid test for CDI were derived based on modelled cost-effectiveness outputs.
- The model analysis comprises of three phases, including: (1) a *de novo* three-stage MPS framework to identify in advance minimum performance specifications based on clinical and cost-effectiveness outputs for each test turnaround time values; (2) univariate deterministic sensitivity and scenario analyses to identify key drivers of cost-effectiveness changes in the model (as identified in phase 1); and (3) univariate deterministic sensitivity and scenario analyses exploring the impact of selected key drivers (identified in phase 2) on the minimum performance specifications identified in phase 1.
- The presented MPS framework represents the first application of EEE methods as a means of deriving minimum performance specifications for key properties of TPPs for tests.

The final chapter (Chapter 7) summarises the key findings and provides a discussion of the thesis.

# Chapter 7
# Discussion

## 7.1 Chapter outline

The final chapter of this thesis discusses the research findings. The main findings from each chapter are first summarised (see section 7.2), followed by an overview of recommendations for developing TPPs (section 7.3), with a focus on the methodology employed to elicit unmet clinical needs (i.e. scoping phase of TPPs) (section 7.3.1) and for integrating EEE methods into TPP development (section 7.3.2). The final sections of this chapter conclude with the limitations of the thesis (section 7.4), suggestions for future research (section 7.5), and an overview of the impact of the presented work (section 7.6).

## 7.2 Research findings

**Chapter 1** of this thesis introduced the use of TPPs as a means of ensuring that innovation and research efforts are focused on tests that are 'fit for purpose'. Based on published examples of early economic models for tests, the argument was made that EEE methods could help to inform key desirable test characteristics for TPPs. The aim of this thesis was therefore to explore how EEE methods could be integrated into the TPP development process for medical tests.

The systematic review presented in **Chapter 2** aimed to identify the current methodology for developing TPPs for medical tests (1). A common development process for TPPs was identified (consisting of scoping, drafting and consensus-building phases), as well as key limitations with the current methods applied within TPPs, including: (i) a heavy reliance on subjective data sources; (ii) poor transparency in reporting the methods underpinning TPP development; (iii) a lack of explicit consideration of clinical utility when defining test specifications; and (iv) an oversight of cost-effectiveness considerations. Based on these limitations, the potential benefits of integrating EEE methods into the development process for TPPs were presented (2). Using EEE in this context was proposed as a means of facilitating a more objective, evidence-based and transparent approach to defining test performance specifications, while also providing a structured approach to understanding and modelling the care pathway within which new tests might sit (2). The ability to iteratively update early economic models as new evidence emerges, and to identify future research priorities based on the findings

of sensitivity and scenario analyses, were highlighted as key advantages of EEEs (2).

The suggested EEE methods were then applied to a case study – a new rapid diagnostic test for CDI (**Chapter 3** to **Chapter 6**) – to explore how EEE methods may be utilised to derive core TPP performance specifications. **Chapter 3** provided the clinical context to the case study – with a particular focus on the laboratory diagnosis of CDI, and the uncertainty around when and who to test for CDI. Two main types of diagnostic strategies for CDI were outlined, including: (i) *standalone tests* – diagnostic tests detecting a single target analyte; and (ii) *testing algorithms* – a combination of two or more standalone tests. The available diagnostic strategies for CDI were presented, followed by an overview of the UK clinical pathway for patients suspected and confirmed with CDI.

To guide the development of a future TPP for new diagnostic tests for CDI, the online survey of UK healthcare professionals presented in **Chapter 4** was conducted as a means of identifying current approaches to CDI diagnosis, unmet clinical needs for new CDI diagnostics, and problems associated with the current diagnostic pathway for patients suspected with CDI. Forty-eight participants completed the survey, the largest group of whom were infection control nurses (n=21), followed by consultants (n=9). Watery stools and unexplained diarrhoea were the main symptoms found to prompt suspicion of CDI among clinicians, with older patients being more frequently suspected of CDI. Due to the lack of an accurate standalone test, a variety of testing algorithms are used across different hospital laboratories, with running EIA GDH and EIA toxin A/B being the most commonly reported testing algorithm. Infection-control measures were usually initiated upon suspicion of CDI, and decisions around when to de-escalate these measures were mostly driven by resolution of symptoms. The primary issues highlighted as hindering the diagnosis of CDI concerned difficulties in identifying who requires testing, and issues with sample collection (i.e. problems in obtaining an adequate stool sample). Test turnaround time was also mentioned as an issue. Key identified requirements for a new CDI diagnostic included: (i) quicker turnaround time; (ii) less invasive sample requirements; and (iii) high diagnostic accuracy.

**Chapter 5** described the structure, development, parameterisation and implementation of a *de novo* early economic model, which aimed to compare the

cost-effectiveness of a new hypothetical point-of-care diagnostic test (HT) for CDI against standard care. Findings from a rapid literature review of published decision models for CDI diagnostics (n=8) informed the development of the structure and parameterisation of the model, while also highlighting key methodological limitations of past models. These included: (i) a failure to capture the full range of different events associated with the clinical pathway for CDI; and (ii) a weak evidence base in relation to the health-related utility weights for CDI. In addition, the majority of past models failed to capture capacity considerations when evaluating the cost-effectiveness of CDI diagnostics. A *de novo* deterministic resource-constrained DES model was then developed to reflect the key processes underpinning the diagnostic and clinical pathways for patients suspected with CDI, based on data from COMBACTE-CDI datasets, literature and consultations with clinical experts in the absence of published data. The model is able to simultaneously capture the movement of patients through the hospital (e.g. single room isolation, stay in general ward) and the testing processes undertaken on each patient's test sample, while accounting for the key capacity constraints of limited single rooms for patient isolation.

**Chapter 6** presents the model analysis, which aimed to identify minimum performance requirements and maximum costs for new rapid tests for CDI, based on cost-effectiveness considerations. The model analysis comprised of three core phases: (1) a novel three-stage MPS framework used to identify minimum performance specifications for HT using the NICE WTP threshold per QALY gained; (2) a deterministic univariate sensitivity analysis and scenario analyses on the clinical and cost-effectiveness outcomes associated with the minimum performance specifications to identify key drivers of change; and (3) a deterministic univariate sensitivity analysis and scenario analyses to explore how the minimum performance specifications identified in phase 1 would change according to changes in the values of key drivers identified in phase 2. The results of the MPS framework and broader model analysis presented in **Chapter 6** illustrate how EEE methods can be utilised to derive minimum key performance requirements for TPPs, and isolate influential drivers of cost-effectiveness considerations. These findings should be discussed with clinical experts and relevant stakeholders to define the minimum and optimal performance requirements for rapid tests for CDI within a future TPP.

## 7.3 Implications of findings and suggested recommendations for developing TPPs for medical tests

The variability in the methods employed to draft TPPs found in the systematic review of TPPs presented in Chapter 2, as well as the poor transparency in methodological reporting, indicate the need to develop guidance for drafting TPPs for tests. The increased interest and use of TPPs for medical tests during the COVID-19 pandemic heightens the need for formal guidance as to best practice methods for producing clinically meaningful, evidence-based TPPs.

Below general considerations for TPP development are summarised, followed by more specific discussion of recommendations for: (i) the scoping phase of TPPs – based on the findings of the systematic review in Chapter 2, and the online survey in Chapter 4 (see section 7.3.1); and (ii) the integration of EEE methods into the TPP development – based on the model findings from Chapter 5 and Chapter 6 (see section 7.3.2).

Based on the findings of Chapter 2, future guidance in this context could feature the following general recommendations for TPP development, including:

- **transparent reporting of the data sources used to inform a TPP** – a minimum level of methodological reporting should be required to draft well-developed TPPs (e.g. providing a biography of literature reviewed, justification for selected data sources, providing details on which clinical experts and stakeholders took part in consultations, listing all activities undertaken to develop a TPP). As well as improving transparency, this would help in the process of updating the TPP once new evidence emerges.
- **systematic assessment of the data sources to inform a TPP** – this guidance document could also provide developers of TPPs with recommendations on: (i) which source of evidence to review first (e.g. literature, modelling studies, available data); and (ii) how to proceed in case more robust evidence is not available (e.g. recommendations on how to systematically elicit and report views of clinical experts and stakeholders).
- **adherence to a consistent terminology of test characteristics listed in TPPs** – having a consistent and clear terminology for the test

specifications underpinning the key evidence domains (e.g. analytical performance, clinical validity) across different TPPs could increase their comparability, while also enhancing their applicability for the industry. This glossary could be based on published frameworks for the test evaluation pathway (21), in addition to the Clinical & Laboratory Standards Institute (CSLI) Harmonized Terminology Database for analytical and technical terminology (321).

### 7.3.1 Recommendations for scoping phase of TPPs

Based on the findings of the systematic review discussed in Chapter 2, as well as the results and limitations of the online survey presented in Chapter 4, this section provides recommendations related to the scoping phase of TPPs, with a focus on: (a) what are the requirements to conduct a well-executed scoping phase for a TPP (section 7.3.1.1); and (b) how to effectively elicit stakeholders' views on unmet clinical needs (see section 7.3.1.2).

#### 7.3.1.1 What are the requirements to conduct a well-executed scoping phase for a TPP?

Core to the scoping phase of a TPP is to engage with a multidisciplinary panel of stakeholders and clinical experts to identify what clinical needs a TPP should address (1). Engagement and recruitment of healthcare professionals into research projects, however, can be challenging due to several barriers. These include: (i) lack of respondents' time and additional project-related work required; (ii) lack of interest in the research project due to poorly formulated or irrelevant study questions, or limited familiarity with the topic being evaluated; and (iii) difficulties in accessing databases of healthcare professionals due to data protection (322, 323). For instance, clinician response rates to surveys are usually low (324, 325), with the response rates from primary care professionals typically ranging between 10.3% to 61% (326-329).

The challenges in engaging with clinical experts were demonstrated in the online survey of UK healthcare professionals presented in Chapter 4. The lower-than-expected uptake and difficulties in the dissemination of the online survey on CDI diagnostics could be due to several factors, including (a) difficulties in recruiting time-limited respondents in the context of the COVID-19 pandemic; (b) lack of engagement from the relevant opinion leaders in the field; or (c) lack of incentives

for the respondents to complete the survey. Healthcare professionals receiving monetary or non-monetary incentives are more likely to fill in a survey compared to professionals given no incentive (324, 330, 331).

The barriers to the dissemination strategy of the online survey on CDI diagnostics suggest key requirements for conducting a well-executed scoping phase for a TPP. These include: (i) having access to a large network from which to draw key thought leaders and respondents; (ii) sufficient time and resources for recruitment (e.g. monetary and non-monetary incentives); and (iii) the possibility to organise focus groups to elicit initial suggestions in regards to clinical problems in the disease of interest, or to validate survey findings. These key requirements indicate that TPP scoping activities are more likely to be successful when significant financial and infrastructure support is available (e.g. with support from a large healthcare organisation or charity, such as FIND or CRUK) to be able to invest in wide-scale recruitment and attract engagement with key opinion leaders and relevant stakeholders.

Alternatively, in the context of limited financial and infrastructure support, TPP scoping activities may still have a high probability of success in scenarios where the disease area being investigated is considered a health priority. Surveys focusing on topics of high interest usually attract more respondents among healthcare professionals (325). For example, an online survey aiming to identify and prioritise use cases for new COVID-19 diagnostics yielded 447 responses despite having a short dissemination period (25 days) (73). Compared to the online survey on CDI presented in Chapter 4, response rates for the COVID-19 survey were substantially higher, presumably as many clinicians regarded the COVID-19 pandemic as a key priority. In addition, at the early stages of the pandemic, testing strategies had a key role in curbing the spread of infection thereby explaining the higher interest for the online survey on COVID-19 diagnostics.

#### 7.3.1.2 How to effectively elicit stakeholders' views on unmet clinical needs?

Based on the findings of the systematic review of TPPs (see Chapter 2), a variety of methods were employed to scope the unmet clinical need of interest. This lack of a consistent methodology to the scoping phase indicates a need for a more structured and effective approach to identifying the clinical problems TPPs should address. Based on the design and validation process of the online survey presented in Chapter 4, key suggestions relevant to the elicitation of unmet clinical needs TPPs should address are discussed below.

Scoping unmet clinical needs might require several sequential activities to come to an agreement on which clinical problems a TPP should focus on – particularly when dealing with a complex testing pathway (such as that for patients suspected of CDI), or 'hard to reach' clinical audiences. In such cases, unless data saturation is reached during the analysis of survey responses, further rounds of dissemination and/or consultations with clinical experts and relevant stakeholders might be required to reach consensus. A possible alternative approach may be to base initial suggestions of potential issues hindering the care pathway of interest on literature findings and initial focused consultations with relevant stakeholders, and later expand on and validate those initial findings with a broader audience (e.g. via an online survey). Based on best practices for eliciting unmet clinical needs of the principal investigators of the National Institute of Biomedical Imaging and Bioengineering Point-of-Care Technologies Research Network (POCTRN), qualitative research methods (e.g. focus groups and literature reviews) are a useful first step to understand the context of interest and to gather perspectives from various stakeholders (5). Results from qualitative studies are then helpful to devise larger and more quantitative surveys which allows the researcher to obtain responses from a wider cohort on a smaller, more focused set of questions (5).

In addition, a key first step in understanding the unmet clinical need and clinical problems in a certain disease area is to map the care pathway in which the new (and existing) technology will sit. Unmet clinical needs are rooted in the care pathway of interest (18), which is dependent on the clinical setting. While the online survey presented in Chapter 4 was initially based on the EFLM checklist for identifying unmet clinical needs for new biomarkers (17, 18), the original

checklist was heavily edited during the design and validation process to make the questions more tailored to the disease of interest and to better map the care pathway being evaluated. Asking respondents to indicate clinical problems affecting a certain disease with no contextual questions on the care pathway of interest was perceived as challenging. The feedback received during the pre-testing phase of the survey suggested that first mapping the care pathway of interest provides a solid foundation for understanding what issues affect that given pathway. Instead of asking generic questions (e.g. “What are the unmet clinical needs in the context of this disease?”), a more efficient approach is to first describe the processes in place for diagnosing a certain condition, and then to identify the issues hindering the diagnosis of that condition. Mapping the care pathway in advance also helps to specify any sources of variation in the testing protocol (e.g. what testing strategies are used, when to test patients) and clinical decision-making across different settings.

While the use of conditional branching questions allowed customisation of the questions being asked as part of the survey presented in Chapter 4, conducting an online survey might be an impractical approach to map the typical care pathway under evaluation – especially where several testing options or sequential steps are included as part of the pathway. In the context of complex care pathways, conducting a focus group of clinical experts with the possibility of asking follow-up questions might have been a more efficient approach to initially determine the care pathway. Key differences within the pathway of interest (e.g. testing strategy option) could then be investigated at later stages via an online survey to identify the most frequent option across different clinical settings.

### 7.3.2 Recommendations for integrating EEE methods into TPP development

This thesis has provided a pragmatic approach to derive minimum performance specifications for new (and existing) tests based on clinical- and cost-effectiveness considerations. The proposed MPS framework and model analysis presented in Chapter 6 represent the first application of EEE methods used to inform TPP performance specifications. This use of EEE represents a major shift in TPP methodology, and is expected to pave the way for similar studies in the future. Guidance as to when and how to integrate EEE methods in the typical TPP development process would therefore be highly beneficial. In particular, future recommendations should address four key questions: (i) *what* is the role of early economic models in the context of TPP development; (ii) *what* are the main requirements for a model-based EEE in the context of TPP development; and (iii) *how* should the MPS framework be applied to derive TPP specifications.

#### 7.3.2.1 What is the role of early economic models in the context of TPP development?

Results from an early economic model are useful to narrow down the range of quantitative performance specifications (e.g. diagnostic accuracy, test price) which result in the test of interest being cost-effective, and also demonstrates how different test properties (e.g. diagnostic sensitivity and specificity) and other factors of the clinical context interact (2). The intended use of EEE results in the context of TPP development is therefore to provide clinical stakeholders and experts with evidence-based data to define minimum and optimal TPP performance requirements based on cost-effectiveness considerations, as opposed to drawing definitive conclusions on the cost-effectiveness of a hypothetical test. Rather than replacing other data sources within TPP development, EEE results in this context should be regarded as a means of informing and supporting a group consensus (2). Because of the inevitable amount of uncertainty within a model-based EEE, consultations with relevant stakeholders should be regarded as key to ensure that the early economic model captures the nuances of the clinical context, and that the estimated performance specifications using the MPS framework are technically and clinically feasible.

Ideally stakeholders and clinical experts should be engaged throughout the model development, so that they understand and agree with the model scope,

structure and assumptions, and can interpret the results in light of these aspects. While the importance of clinical input and feedback is relevant for every model, in the context of TPP development stakeholder input is pivotal to ensuring that the application of the MPS framework is 'fit for purpose', and that the modelling results are clinically and economically valid. Details on the key aspects TPP developers and relevant stakeholders should discuss prior to develop an early economic model and to run the MPS framework are presented in section 7.3.2.3.

#### 7.3.2.2 What are the main requirements for a model-based EEE in the context of TPP development?

Based on the strengths and limitations of the model presented in Chapter 5, key requirements are suggested for developing an early economic model as a means of informing TPP specifications. The following sections discuss each of these requirements.

**Modelling technique should be 'fit for purpose'**

Initial discussion should take place at the early TPP development stages as to which modelling technique to adopt. As discussed in Chapter 5, this choice depends on the decision being evaluated, as well as which key mechanisms must be captured within the model. An additional consideration is the intended value proposition of the hypothetical test being evaluated, as well as which performance dimensions to capture within the model. For example, if there was a need for a new rapid test, modelling techniques where the timing of events are captured would be most suitable for deriving minimum specifications for test turnaround time.

Depending on the research question at hand and the test performance dimensions being evaluated, simpler decision models (e.g. decision trees or Markov models) could be used to derive minimum performance specifications. For example, decision trees are often used to identify minimum diagnostic accuracy and maximum price for a hypothetical test based on cost-effectiveness considerations – as discussed in Chapter 1. The crucial point is that, whatever modelling approach is adopted, the processes underpinning the model development and selection of model data should be fully transparent and justified.

**Balance granularity and generalisability of the model structure**

Modellers should decide at early TPP development stages as to whether the model structure should resemble clinical practice in a specific setting or be generalisable across different clinical settings. If there are significant regional differences in the clinical pathway being evaluated, it is suggested to focus on a specific clinical setting (e.g. LTHT as in the model outlined in Chapter 5), while also allowing for the possibility of exploring the impact of different clinical practices and assumptions via scenario analyses. Consultations with clinical experts should inform the most relevant clinical scenarios to explore as part of the model analysis, as presented in Chapter 6 (i.e. phase 2). This should help to focus the model on clinically relevant scenarios, reducing the risk of overwhelming stakeholders with excessive information.

**Adopt a technology-agnostic perspective when appropriate**

The degree of detail surrounding the outline of the hypothetical test being evaluated (e.g. number of samples being tested, how many testing kits are available in a certain clinical setting) should reflect the stage of the TPP being developed. Technology-agnostic early economic models evaluating a generic hypothetical test would better support the development of early TPPs aiming to defining the key requirements for new tests to address identified unmet clinical needs. In this context, it is recommended against making unnecessary assumptions about the hypothetical test – unless these assumptions are supported by consultations with clinical experts. For example, modelling techniques with the ability to capture key processes underpinning the testing workflow (e.g. DES and ABS models) allow modellers to simulate in granular detail the main features of the hypothetical test which, without the necessary clinical input, might appear arbitrary and redundant. Later model iterations could capture more detailed information and assumptions on the test being evaluated once new evidence on the desirable performance specifications and the clinical pathway emerges. More detailed TPPs may also be produced at later stages when specific technologies are starting to be developed and to enter the market (i.e. moving beyond the purely ‘hypothetical’ stage).

**Develop a flexible, transparent and open-access model**

The developed model should be sufficiently flexible to conduct extensive sensitivity analyses without requiring substantial time to run multiple simulations. Developing a flexible model also allows simulation of the impact of alternative clinical scenarios on the model outcomes via scenario analyses, and ensures the model structure and parameters can be easily updated once new evidence emerges.

In addition, early economic models in this context should not be perceived as 'black boxes' by the clinical stakeholders. Instead, full model documentation should be made available providing accessible information on: (i) clinical assumptions; (ii) structure; (iii) parameterisation; (iv) validation; and (v) analysis. Ideally, model-based EEEs in this context should also be programmed in free/widely available software (e.g. Microsoft Excel, R, Python) and the model coding made publicly available in open access platforms (e.g. GitHub). This also allows the model parameters and structure to be updated once new evidence becomes available, thereby ensuring that TPP performance specifications are efficiently updated over time.

#### 7.3.2.3 How should MPS framework be applied to derive TPP specifications ?

While the framework developed in this thesis focused on a specific case study, the proposed methods could be extended to different clinical areas. To ensure that the proposed MPS framework is 'fit for purpose' in the context of TPP development and to make it more generalisable to other clinical settings, initial discussions among TPP developers and relevant stakeholders should take place around: (1) what test performance dimensions to capture within the early economic model; (2) what is the meaning of 'minimum' and 'optimal' test requirements; and (3) what model outputs should be considered to derive minimum performance requirements for HT.

Each of these key questions are discussed below, based on the model analysis and MPS framework presented in Chapter 6.

**What test performance dimensions should be included within the model?**

While defining the minimum and ideal properties for a hypothetical test, high uncertainty is unavoidable as test specifications may vary when considering

different aspects underlying either the test itself or the care pathway (2). Initial discussions with clinical experts and relevant stakeholders are therefore an important step for determining what key minimum performance requirements should be derived using the MPS framework. This pragmatic decision should reflect the unmet clinical need for tests in the context of the disease of interest, and most importantly, the expected value proposition of the hypothetical test. For example, if a need for a rapid and accurate test emerges during the scoping phase, the model should capture both diagnostic accuracy and test turnaround time as key performance dimensions.

The early economic model presented in Chapter 5 and Chapter 6 focused on minimum performance specifications for diagnostic accuracy, test turnaround time and cost, but it is plausible that the MPS framework could be expanded to inform additional aspects of the TPP – such as *analytical performance requirements* (e.g. maximum bias and imprecision of the test). Past authors have used decision models to show the link between changing analytical validity (e.g. bias and imprecision) and clinical outcomes, indicating that it is possible to integrate elements of analytical validity into decision models evaluating the cost-effectiveness of alternative testing strategies (332, 333). This would, however, increase the complexity of the model structure and parameterisation. In particular, the modeler would need data to inform how analytical validity influences the diagnostic accuracy of the test under consideration. Data on this, however, would most likely be unavailable for a hypothetical test. Whilst there may not be sufficient data to model the impact of analytical validity on outcomes in the context of hypothetical tests, once the clinical performance benchmarks for a *'real'* test are derived, these could be used by laboratory professionals to help them derive maximum allowable levels of bias and imprecision.

Clearly as the number of performance values to be addressed increases, the complexity of the analyses required also rapidly expands as well as the risk of overwhelming relevant stakeholders with excessive modelling results. This, in turn, indicates that the decision around which test performance dimensions to include in the model should also reflect practical factors, such as model computational time and project deadlines. In the context of time-constrained projects, a viable approach would be to predominantly focus on test characteristics which drive the unmet clinical need (e.g. need for a more accurate

test, or need for a quicker test result). Deriving performance benchmarks for analytical validity within this framework is possible, although the methods to do this are not currently part of routine decision modelling practice. This requires further research and additional case studies to better understand how to integrate these new methods into the TPP development process.

**What is the definition for 'minimum' and 'optimal' test requirements?**

Based on the findings of the systematic review presented in Chapter 2, there is no common definition for 'minimum' and 'optimal' requirements applied across TPPs – coupled with a poor transparency in providing a definition for these key terms (1). This suggests the need to adhere to a consistent and transparent terminology throughout the TPP development process.

Initial discussion within clinical experts and relevant stakeholders should therefore take place in the early TPP development phases, to agree on the definitions for 'minimum' and 'optimal' requirements. Based on past TPPs, 'optimal' requirements for diagnostic accuracy were often set at over 99% for sensitivity and/or specificity, without any clear assessment of whether such optimal requirements could be feasibly achievable by test manufacturers. As diagnostic accuracy is bounded between 0% and 100%, the theoretical 'optimal' target will always be perfect diagnostic accuracy (100% sensitivity and specificity). This, however, ignores: (i) what new (or existing) testing technologies can feasibly achieve; and (ii) the current accuracy of the diagnostic reference standard. For instance, assuming that the standard testing strategy for a certain disease has a diagnostic sensitivity of 80%, setting an optimal requirement of 100% for diagnostic sensitivity for any new competitor diagnostics is likely to be unrealistic for test manufacturers and may even stifle innovation or encourage biased study designs and reporting. Initial consultations with clinical experts and relevant stakeholders are therefore pivotal to set clear definitions for 'minimum' and 'optimal' requirements to ensure that an agreed and consistent definition is applied throughout the TPP development process. In this context, future qualitative research could investigate what key factors stakeholders consider when defining 'minimum' and 'optimal' requirements (e.g. performance and characteristics of the diagnostic reference testing and/or available testing options).

**What model outputs should be considered to derive the minimum performance specifications for HT?**

To make the MPS framework 'fit for purpose' in the context of TPP development, a first task for TPP developers and relevant stakeholders is to reach a consensus on the jurisdiction, targeted audience and primary outputs of the early economic evaluation being conducted. Parameterisation and primary outcomes of economic evaluations – early or otherwise – are dependent on the jurisdiction and decision-maker of interest who, in turn, dictates what pre-specified decision criterion is used to inform the analysis. The MPS framework is expected to produce meaningful minimum performance specifications providing that it evaluates output(s) and applies a decision criterion which is ultimately relevant for the decision-maker(s) of interest. For example, were NICE the targeted audience of the early economic evaluation being conducted, the MPS framework should apply the WTP per QALY gains to estimate the minimum performance specifications. At early stages of TPP development, however, it may be plausible that there is no specific decision-maker for the economic evaluation being conducted given the early stages of product development and technology assessment. It is therefore recommended to consider different model outputs that might be of interest to the decision-maker(s) in charge of the procurement process of new tests in a given jurisdiction (e.g. NICE and NHS commissioners in the UK). This, in turn, is expected to increase the perceived utility of the MPS framework and the applicability of the estimated minimum performance specifications.

In addition, it is important to liaise with clinical experts and stakeholders to address four key questions relating to what model outputs to use to derive minimum performance specifications. These questions include: (i) were the QALY gains deemed to be an output of interest and do they sufficiently capture the (expected) benefits of introducing a new diagnostic test on patients' health and the clinical pathway of interest; (ii) what intermediate clinical model outputs should be considered when evaluating the clinical effectiveness of HT; (iii) in the context of multiple clinical outputs of interest, how should these outputs be combined to define when the HT is considered clinical inferior/superior compared to standard care; and (iv) whether it is appropriate to exclude clinical inferior

options which may be considered cost-effective due to cost savings. Each of these key questions are discussed below.

**The use of WTP threshold per QALY gains: do QALY gains fully capture the expected benefits of a new test?**

Core to the presented MPS framework is the use of a pre-specified WTP threshold to derive the minimum performance requirements for a hypothetical test to maintain cost-effectiveness compared to standard care. The suggested EEE methods and MPS framework using the WTP threshold per QALY gains are expected to provide informative results (i.e. minimum performance specifications) in the context of health conditions where at least one of the following requirements is met:

- test results inform clinical decision-making regarding treatment – i.e. laboratory results play a major role in deciding whether or not to administer treatment to patients;
- there is evidence (or a justification, in the context of a paucity of data) that the test-informed treatment decisions will have an impact on patients' survival and/or quality of life

Were one of these requirements be met, the proposed EEE methods and MPS framework would provide meaningful minimum performance specifications for a test to maintain cost-effectiveness as there is sufficient evidence (and/or a justification) that a test could have an impact on patients' survival and/or quality of life. Alternatively, in a scenario where there is limited data on the impact of a test on these outcomes, or where several assumptions need to be made to capture the impact of testing on patients' long-term health outcomes, the value of deriving the minimum performance specifications using the WTP per QALY gains might be questionable. Difficulties in capturing the impact of testing strategies on QALYs, however, is a common issue among economic evaluations for diagnostic tests given the indirect impact tests have on patient health which is usually mediated through treatment effectiveness (21). Even in the context of a case study where there is (currently) a paucity of evidence and/or justification of the impact of the test-treatment pathway on QALY gains, future iterations of the model could include additional data when made available, with the possibility of re-running the MPS framework to derive up-to-date performance requirements.

In the context of limited evidence (and/or justification) of the impact of the test-treatment pathway on patients' survival and/or quality of life, or in the context of other outcomes of interest to relevant decision-maker(s), alternative clinical outcomes could be measured to quantify the full impact a new test, beyond the QALY. Assigning an explicit monetary value to secondary clinical outcomes, however, is challenging, as an explicit WTP threshold is missing in this case – outside of the assessment of QALY gains. A possible alternative, in cases where a different clinical outcome is of primary importance, could be conducting a WTP elicitation exercise with the clinical stakeholders involved in the TPP development process, or additional participants where necessary. This would allow the methodology as presented in Chapter 6 to be employed, using the elicited WTP value as a means of calculating acceptable test performance levels based on cost-effectiveness considerations. The trade-off in this case would be the additional time and resources required to run such an elicitation exercise.

**What clinical model outputs should be considered within the MPS framework?**

In the first stage of the MPS framework (i.e. *assessment of clinical -effectiveness outputs*), the aim was to rule out performance specifications where HT would be considered clinically inferior compared to standard care based on selected clinical outputs. Within the model analysis outlined in Chapter 6, HT was considered to be clinically inferior compared to standard care when: (i) HT led to negative incremental QALYs, and (ii) HT increased the number of secondary infections. As this model aimed to explore the impact of a new POCT on reducing the infection spread, the number of new secondary infections prevented was regarded as an appropriate metric of clinical effectiveness for HT – in addition to QALY gains. This decision, however, was not supported by any discussion with clinical experts or key stakeholders.

In case studies where intermediate clinical outputs are considered of importance by the decision-maker(s) of interest, initial discussions with clinical experts should focus on what intermediate clinical model outputs to consider when evaluating the clinical effectiveness of HT (e.g. number of secondary infections averted, reduction in inappropriate antimicrobial prescriptions) – in addition or as a substitute of the QALY gains. The definition of when a HT is considered clinical inferior/superior, compared to standard care should be dependent on clinical

model outputs that are ultimately relevant for stakeholders and decision-makers of interest. In addition, in the context of multiple clinical outputs of interest, it is important to reach an agreement on how multiple clinical outputs should be combined in order to define when the HT is considered clinical inferior/superior compared to standard care. This, in turn, will help ensure that the MPS framework produces results that are considered clinically meaningful to stakeholders, in line with a consensus-based definition of clinical effectiveness.

**Is it appropriate to exclude clinical inferior options?**

An additional aspect to discuss with clinical experts and relevant stakeholders relates to the relative importance of cost-savings over clinical-effectiveness when determining if HT is cost-effective versus standard care. In the second stage of the MPS framework (i.e. *assessment of cost-effectiveness outputs*) presented in Chapter 6, the aim was to rule out combinations of diagnostic accuracy where HT was not expected to be clinically- and cost-effective – based on the selected clinical outputs. Healthcare interventions with lower clinical benefits but leading to cost savings, however, would still be considered cost-effective if they yield a positive INMB (i.e. if the cost savings are considered sufficient to outweigh the health losses) (61).

In the context of the case study, the exclusion of clinically inferior options was considered appropriate given that the hypothetical POCT was not expected to deliver significant cost savings compared to standard care testing practice – especially since the latter involves batch testing. In addition, clinically inferior testing strategies would not be expected to be acceptable in the context of TPP development. In other clinical contexts, however, cost savings could be of equal or more interest to relevant stakeholders; it is important therefore that initial discussion should take place to ensure that the relevant clinical and economic outcomes, as well as their relative importance, are effectively captured within the MPS framework. For instance, relevant stakeholders and decision-makers might be interested in a new test which (i) improves patient health outcomes but it leads to higher costs; or (ii) is both more clinically effective and cost saving compared to standard care. These decisions reflect the value proposition of a new test and ultimately the unmet clinical need the new test and the TPP are attempting to fulfil. Stating upfront the relative importance of clinical effectiveness and cost-

saving, and how this stems from the value proportion of a new test, is expected to increase the interpretability and generalisability of the MPS results.

Addressing key discussion points presented in this section will ensure that the MPS framework produces evidence-based minimum performance requirements that are clinically and economically meaningful to the relevant stakeholders involved in the TPP development process.

## 7.4 Thesis limitations

Limitations of the thesis have been discussed in depth within each chapter. Key overarching limitations of the thesis are summarised here. A key limitation of this thesis lies in the limited engagement with clinical experts and relevant stakeholders throughout different stages of the research. As mentioned in the COVID-19 impact statement, it was not possible to establish a multidisciplinary stakeholder group to present and discuss the methodology and findings of the presented research. This limitation affected different stages of the presented work, including:

- **online survey presented in Chapter 4** – the low response rate, coupled with a lack of engagement with key stakeholders in the field of CDI diagnosis (e.g. patient and industry representatives), might limit the generalisability of the survey findings. Had patients' preferences been captured in the survey, findings would be expected to indicate that most perceived problems hindering the diagnosis of CDI revolved around issues patients directly experience, such as the sampling (e.g. patients' acceptability to sampling) and post-analytical phase of testing (e.g. how test results are communicated). This, in turn, might have changed the selection of the test characteristics to be investigated as part of the model presented in Chapter 6. To address this limitation, further rounds of dissemination, or alternatively presenting the results to a multi-disciplinary stakeholder panel, would help to validate the survey findings with different respondent groups from those targeted as part of the dissemination strategy.
- **MPS framework and model analysis presented in Chapter 6** – while the structure of the model was developed based on several consultations with clinical experts (see Chapter 5), many decisions around the implementation of the MPS framework were not discussed with or

presented to relevant stakeholders due to the COVID-19 pandemic. For example, key decisions around: (1) what performance dimensions to capture within the model; (2) the meaning of 'minimum' and 'optimal' test requirements; and (3) what model outputs should be considered to derive minimum performance specifications for HT (see section 7.3.2.3) were not presented and discussed with a TPP development expert stakeholder group due to difficulties in engaging with stakeholders in this context because of the COVID-19 pandemic. It is therefore unclear at this stage if the expert stakeholder would have trusted presented modelling results sufficiently to guide the TPP performance specifications. Consultations are expected in the future, as part of the wider research project (see Chapter 1 for more detail), in which results from the MPS framework will be discussed with relevant stakeholders and clinical experts, and used to agree on the final performance specifications for a future TPP on new diagnostics for CDI.

## 7.5 Future research recommendations

Based on the presented work, key areas for future research are recommended. As discussed in Chapter 1, historically there have been three applications of TPPs for healthcare products: (i) in-house industry-led documents; (ii) voluntary briefing documents; and (iii) guiding documents to stimulate innovation (46, 50). Although the scope of this thesis was on TPPs in general in the context of medical tests, the research presented in this thesis focused on the third application of TPPs – i.e. as guiding documents to steer and support the development process of new 'fit for purpose' healthcare technologies in response to perceived unmet clinical needs. This might be explained by the records retrieved and included as part of the systematic review of TPPs presented. As discussed in Chapter 2, the publicly available TPPs for tests were primarily developed by global health organisations as a guiding document to identify necessary features of new tests to address a health priority. Despite this, the methodological suggestions discussed in this thesis are expected to be useful across all three contexts given the similarities across the different applications of TPPs (i.e. the concept of 'beginning with the goal in mind', and the concept of 'living' documents). More (publicly available) applications, however, are needed to confirm this hypothesis. Future research should therefore investigate how industry-led and regulatory

TPPs are currently being developed (e.g. typical decision-making process, input sources, stakeholders involved in the TPP development), and to explore the utility of the presented MPS framework and EEE methods in those settings.

More research is also required to understand whether and how to include other test characteristics (e.g. analytical validity, human factors, infrastructural requirements) into the proposed MPS framework. Focusing on analytical performance requirements, a methodological framework for deriving outcome-based analytical performance specifications based on modelled clinical utility and cost-effectiveness outcomes has previously been developed (333). Similarly, a conceptual framework called *Medical Device Usability Stressor-based Assessment* (MEDUSA) was developed to integrate human factors (e.g. persons, technological context, tasks performed, organisation, physical environment) into early economic models of medical tests (334). Future research could therefore attempt to integrate these methods into the proposed MPS framework as a means of informing other test characteristics based on cost-effectiveness considerations.

Future work should also focus on how to best present modelling results and reach consensus-based performance specifications. A potential way forward to efficiently present results from the MPS framework and the extensive sensitivity analyses could be to use interactive web browser-based user interfaces, such as R Shiny apps. Stakeholders and clinical experts could use these web apps to independently vary the performance dimensions simulated within the model, and assess the impact of varying model parameters and structural scenarios on the minimum performance specifications identified using the MPS framework – without the need to run or understand the model coding. In addition, R Shiny apps uses graphical interfaces to efficiently present modelling results which could support discussion of the MPS framework results among clinical experts (335). R Shiny apps could be useful during the TPP development process as their application has been recommended especially in the context of: (i) high uncertainty surrounding model parameters; (ii) multiple clinical assumptions to test; (iii) the audience of the model comprises of multiple stakeholders; and (iv) model structure and parameterisation need to be quickly updated once new evidence emerges (335).

In addition, future research should explore how to better integrate patient preferences for new tests into the TPP development process. To this end, a Patient and Public Involvement & Engagement (PPIE) group could be established at early stages to ensure that the methods and findings of each phase of TPP development (i.e. scoping, drafting and consensus-building) are communicated, discussed and validated with patients and members of the public. In the context of the scoping phase, initial suggestions on the issues hindering the care pathway of interest – elicited either via a survey or a series of focus groups of healthcare professionals – should always be validated with patient representatives to ensure that patients' preferences are efficiently captured into the definition of the unmet clinical need to address within the TPP. In addition, *discrete choice experiment* (DCE) is a viable means of estimating end-users' preferences towards different test characteristics (e.g. speed of diagnosis, confidence in test results, acceptability of sampling) and health-related outcomes. For example, a recent cross-European DCE study was conducted to explore patients' preferences towards key attributes for diagnostic tests aiming to reduce antimicrobial resistance (336). Similar studies would ensure that new (or existing) tests meeting TPP performance specifications would also accommodate patients' preferences. The establishment of a PPIE group and DCE studies, in turn, would positively influence the development of TPPs and potentially maximise the uptake and adoption of these tests into clinical practice.

Furthermore, future TPP documents should better indicate what evidence and studies are required for manufacturers to prove that their test fulfils the listed desirable criteria. Having an explicit link between the desirable performance targets and what evidence is required to fulfil these criteria could provide the basis for a more efficient evidence-generation pipeline. This may be particularly relevant in the context of TPPs used by regulatory bodies. For example, if national regulators are going to demand certain (minimum) levels of performance, guidance on the study design/research required to obtain evidence on these performance requirements would be beneficial for test developers.

Finally, it would also be valuable to explore if and how TPPs should be integrated into existing regulatory paths for innovation such as the EU IVDR (Regulation 2017/746). It might then be possible to align test characteristics featured in TPPs with evidence requirements that are relevant for market approval decisions of

new medical tests. Future research should assess if a regulatory application of TPPs in this context could ultimately stimulate and accelerate innovation of clinically useful diagnostic tests and their adoption into clinical practice.

## 7.6 Thesis impact

There has been increased attention on diagnostic innovation and TPPs in recent years and, consequently, findings from this thesis have been of interest to national health organisations, charities, and research groups across the UK – especially in the context of the COVID-19 pandemic. For example, the systematic review presented in Chapter 2 informed the performance items included in the first draft of the MHRA TPPs for rapid COVID-19 diagnostics (162). Additionally, the structure and question wording of the online survey on the use cases for new COVID-19 diagnostic tests (337) was adapted from the online survey for CDI diagnostics presented in Chapter 4. Results from the COVID-19 survey fed into recent UK guidelines around testing for COVID-19 and the NICE early economic model on rapid COVID-19 diagnostics (164).

More broadly, CRUK have added the generation of TPPs for novel ED&D tests as an action within their recently published “Roadmap to the future” (45). Since then, the candidate (Cocco P) has been invited to contribute to the development of a TPP of a new diagnostic test ovarian cancer as part of an international collaboration, funded by the CRUK CanTest Collaborative. An international stakeholder expert group has been established and EEE methods will be used to help inform TPP test characteristics. This will be the first TPP developed in response to the roadmap and will set the precedent methodology-wise for future TPP development in cancer diagnostics.

## 7.7 Thesis summary

The following points summarise the key messages of this thesis:

- Developing TPPs for medical tests is a multidisciplinary iterative process which involves three phases (e.g. scoping, drafting, consensus-building). Literature and consultations with experts and stakeholders have been the main data sources used within TPPs to date. Given current variability in the methods employed to draft TPPs and the poor transparency in methodological reporting, further guidance is required for producing clinically meaningful, evidence-based TPPs.
- Using EEE methods to inform TPPs will provide a more objective, evidence-based and transparent approach to defining combinations of test performance specifications based on cost-effectiveness considerations, while also capturing key aspects underpinning the care pathway of interest. EEE results therefore support the expert group in reaching an evidence-based consensus with regards to test performance specifications. EEE methods are also sufficiently flexible to allow for the model to be iteratively updated once new evidence emerges, while also highlighting key areas for future research.
- A *de novo* approach was developed to identify the minimum performance requirements and maximum costs for new tests, based on cost-effectiveness considerations, while also isolating influential drivers of cost-effectiveness changes. The model analysis comprises of three phases, including: (1) a *de novo* three-stage MPS framework to identify in advance minimum performance specifications based on cost-effectiveness outputs; (2) sensitivity analyses to identify key drivers of cost-effectiveness changes in the model; and (3) sensitivity analyses exploring the impact of selected key drivers on the minimum performance specifications identified in phase 1.
- The added value of this framework lies in formalising and structuring existing EEE methods in a coherent way as a means of informing minimum performance specifications using the WTP threshold as a decision criterion, while also accounting for the (inevitable) uncertainty surrounding the minimum performance requirements. Further guidance, however, is required to ensure that a meaningful integration of EEE methods and results into the typical development process of TPP for medical tests is undertaken.

# References


1. Cocco, P., Ayaz-Shah, A., Messenger, M.P., West, R.M. and Shinkins, B. Target Product Profiles for medical tests: a systematic review of current methods. *BMC Med.* 2020, **18**(1), p.119.
2. Cocco, P., Messenger, M.P., Smith, A.F., West, R.M. and Shinkins, B. Integrating Early Economic Evaluation into Target Product Profile development for medical tests: advantages and potential applications. *Int J Technol Assess Health Care.* 2021, **37**(1), p.e68.
3. Seyhan, A.A. Lost in translation: the valley of death across preclinical and clinical divide – identification of problems and overcoming obstacles. *Transl Med Commun.* 2019, **4**(1), p.18.
4. Price, C.P. and St John, A. Translational health economics—delivering the return on investment for laboratory medicine. *J Lab Precis Med.* 2019.
5. Weigl, B.H., Gaydos, C.A., Kost, G., Beyette, F.R.J., Sabourin, S., Rompalo, A., de Los Santos, T., McMullan, J.T. and Haller, J. The Value of Clinical Needs Assessments for Point-of-Care Diagnostics. *Point Care.* 2012, **11**(2), pp.108-113.
6. Wang, P. and Kricka, L. Current and emerging trends in point-of-care technology and strategies for clinical validation and implementation. *Clin Chem.* 2018, **64**(10), pp.1439-1452.
7. National Cancer Institute Dictionary of Cancer Terms. Definition of biomarker [Online]. [Accessed 25/07/2022]. Available from: https://www.cancer.gov/publications/dictionaries/cancer-terms/def/biomarker
8. Walter, F.M., Thompson, M.J., Wellwood, I., Abel, G.A., Hamilton, W., Johnson, M., Lyratzopoulos, G., Messenger, M.P., Neal, R.D., Rubin, G., Singh, H., Spencer, A., Sutton, S., Vedsted, P. and Emery, J.D. Evaluating diagnostic strategies for early detection of cancer: the CanTest framework. *BMC Cancer.* 2019, **19**(1), p.586.
9. Füzéry, A.K., Levin, J., Chan, M.M. and Chan, D.W. Translation of proteomic biomarkers into FDA approved cancer diagnostics: issues and challenges. *Clin Proteom.* 2013, **10**(1), p.13.
10. François, R., Carmen, L., Yves, L. and Yves, G. Measuring the chronology of the translational process of molecular genetic discoveries. *Clin Chem Lab Med.* 2019, **57**(8), pp.1136-1141.
11. Drucker, E. and Krapfenbauer, K. Pitfalls and limitations in translation from biomarker discovery to clinical utility in predictive and personalised medicine. *EPMA J.* 2013, **4**(1), p.7.
12. Kern, S. Why Your New Cancer Biomarker May Never Work: Recurrent Patterns and Remarkable Diversity in Biomarker Failures. *Cancer Res.* 2012, **72**(23), pp.6097-6101.
13. Thompson, M., Price, C., Van den Bruel, A., Heneghan, C., Wilson, P., Crabb, N., Williams, D.A., Deeks, J. and P.;, B. Innovation in Diagnostics and Healthcare: Improving “bench to bedside” processes for diagnostic testing. [Online]. 2011. [Accessed 20/02/2019]. Available from: https://www.community.healthcare.mic.nihr.ac.uk/reports-and-resources/innovation-in-diagnostics-in-healthcare
14. Engel, N., Wachter, K., Pai, M., Gallarda, J., Boehme, C., Celentano, I. and Weintraub, R. Addressing the challenges of diagnostics demand and

supply: insights from an online global health discussion platform. *BMJ Glob Health.* 2016, (2059-7908 (Print)).
15. Korte, B.J., Rompalo, A., Manabe, Y.C. and Gaydos, C.A. Overcoming Challenges With the Adoption of Point-of-Care Testing: From Technology Push and Clinical Needs to Value Propositions. *Point Care.* 2020, **19**(3).
16. Engel, N., Wachter, K., Pai, M., Gallarda, J., Boehme, C., Celentano, I. and Weintraub, R. Addressing the challenges of diagnostics demand and supply: insights from an online global health discussion platform. *BMJ Glob Health.* 2016, **1**(4), p.e000132.
17. Monaghan, P.J., Robinson, S., Rajdl, D.l., Bossuyt, P.M.M., Sandberg, S., St John, A., O'Kane, M., Lennartz, L., Röddiger, R., Lord, S.J., Cobbaert, C.M. and Horvath, A.R. Practical guide for identifying unmet clinical needs for biomarkers. *EJIFCC.* 2018, **29**(2), pp.129-137.
18. Monaghan, P., Lord, S., St John, A., Sandberg, S., Cobbaert, C., Lennartz, L., Verhagen-Kamerbeek, W., Ebert, C., Bossuyt, P. and Horvath, A. Biomarker development targeting unmet clinical needs. *Clin Chim Acta.* 2016, **460**, pp.211-219.
19. Renee Ruhaak, L., van der Laarse, A. and Cobbaert, C.M. Apolipoprotein profiling as a personalized approach to the diagnosis and treatment of dyslipidaemia. *Ann Clin Biochem.* 2019, **56**(3), pp.338-356.
20. van Duijl, T.T., Soonawala, D., de Fijter, J.W., Ruhaak, L.R. and Cobbaert, C.M. Rational selection of a biomarker panel targeting unmet clinical needs in kidney injury. *Clin Proteom.* 2021, **18**(1), p.10.
21. Horvath, A., Lord, S., St John, A., Sandberg, S., Cobbaert, C., Lorenz, S., Monaghan, P., Verhagen-Kamerbeek, W., Ebert, C. and Bossuyt, P. From biomarkers to medical tests: The changing landscape of test evaluation. *Clin Chim Acta.* 2014, **427**, pp.49-57.
22. Lijmer, J.G., Leeflang, M. and Bossuyt, P.M.M. Proposals for a Phased Evaluation of Medical Tests. *Med Decis Making.* 2009, **29**(5), pp.E13-E21.
23. Burke, W. and Zimmern, R. Moving beyond ACCE: an expanded framework for genetic test evaluation. [Online]. 2007. [Accessed 02/02/2019]. Available from: http://www.phgfoundation.org/documents/369_1409657043.pdf
24. Thompson, M., Weigl, B., Fitzpatrick, A. and Ide, N. More Than Just Accuracy: A Novel Method to Incorporate Multiple Test Attributes in Evaluating Diagnostic Tests Including Point of Care Tests. *IEEE J Transl Eng Health Med.* 2016, **4**, pp.1-8.
25. Bossuyt, P.M., Irwig, L., Craig, J. and Glasziou, P. Comparative accuracy: assessing new tests against existing diagnostic pathways. *BMJ.* 2006, **332**(7549), p.1089.
26. Medical Dictionary for the Health Professions and Nursing. "analyte". [Online]. 2012. [Accessed 25/07/2022]. Available from: https://medical-dictionary.thefreedictionary.com/analyte
27. EUR-Lex. Regulation (EU) 2017/746 of the European Parliament and of the Council of 5 April 2017 on in vitro diagnostic medical devices and repealing Directive 98/79/EC and Commission Decision 2010/227/EU (Text with EEA relevance). [Online]. 2017. [Accessed 02/02/2019]. Available from: https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32017R0746&from=EN
28. Simundic, A.-M. Measures of Diagnostic Accuracy: Basic Definitions. *Med Biol Sci.* 2008, **22**.

29. Abel, L., Shinkins, B., Smith, A., Sutton, A., Sagoo, G., Uchegbu, I., Allen, A., Graziadio, S., Moloney, E., Yang, Y. and Hall, P. Early Economic Evaluation of Diagnostic Technologies: Experiences of the NIHR Diagnostic Evidence Co-operatives. *Med Decis Making.* 2019, **39**(7), pp.857-866.
30. Bossuyt, P., Reitsma, J., Linnet, K. and Moons, K. Beyond diagnostic accuracy: the clinical utility of diagnostic tests. *Clin Chem.* 2012, **58**(12), pp.1636-1643.
31. Verbakel, J., Turner, P., Thompson, M., Plüddemann, A., Price, C., Shinkins, B. and Van den Bruel, A. Common evidence gaps in point-of-care diagnostic test evaluation: a review of horizon scan reports. *BMJ Open.* 2017, **7**(9), pp.1-9.
32. Borsci, S., Buckle, P. and Hanna, G.B. Why you need to include human factors in clinical and empirical studies of in vitro point of care devices? Review and future perspectives. *Expert Rev Med Devices.* 2016, **13**(4), pp.405-416.
33. EUR-Lex. Directive 98/79/EC of the European Parliament and of the Council of 27 October 1998 on in vitro diagnostic medical devices. [Online]. 1998. [Accessed 01/04/2022]. Available from: https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:31998L0079&from=EN
34. European Commision. Public health: Commission proposes a progressive roll-out of the new In Vitro Diagnostic Medical Devices Regulation. [Online]. 2021. [Accessed 31/03/2022]. Available from: https://ec.europa.eu/commission/presscorner/detail/en/IP_21_5209
35. Cobbaert, C., Capoluongo, E.D., Vanstapel, F.J.L.A., Bossuyt, P.M.M., Bhattoa, H.P., Nissen, P.H., Orth, M., Streichert, T., Young, I.S., Macintyre, E., Fraser, A.G. and Neumaier, M. Implementation of the new EU IVD regulation – urgent initiatives are needed to avert impending crisis. *Clin Chem Lab Med.* 2022, **60**(1), pp.33-43.
36. Soares, M.O., Walker, S., Palmer, S.J. and Sculpher, M.J. Establishing the Value of Diagnostic and Prognostic Tests in Health Technology Assessment. *Med Decis Making.* 2018, **38**(4), pp.495-508.
37. World Health Organisation Europe. Health technology assessment. [Online]. [Accessed 02/04/2022]. Available from: https://www.euro.who.int/en/health-topics/Health-systems/health-technologies-and-medicines/policy-areas/health-technology-assessment
38. Cowles, E., Marsden, G., Cole, A. and Devlin, N. A Review of NICE Methods and Processes Across Health Technology Assessment Programmes: Why the Differences and What is the Impact? *Appl Health Econ Health Policy.* 2017, **15**(4), pp.469-477.
39. Llewellyn, S., Procter, R., Harvey, G., Maniatopoulos, G. and Boyd, A. Facilitating technology adoption in the NHS: negotiating the organisational and policy context – a qualitative study. *Health Serv Res.* 2014, **2**, pp.1-132.
40. Maniatopoulos, G., Haining, S., Allen, J. and Wilkes, S. Negotiating commissioning pathways for the successful implementation of innovative health technology in primary care. *BMC Health Serv Res.* 2019, **19**(1), p.648.
41. Ivanov, A. Barriers to the Introduction of New Medical Diagnostic Tests. *Lab Med.* 2013, **44**(4), pp.e132-e136.

42. Billings, P.R. Three barriers to innovative diagnostics. *Nat Biotechnol.* 2006, **24**(8), pp.917-918.
43. Eisenberg, J.M., Schwartz, J.S., McCaslin, F.C., Kaufman, R., Glick, H. and Kroch, E. Substituting Diagnostic Services: New Tests Only Partly Replace Older Ones. *JAMA.* 1989, **262**(9), pp.1196-1200.
44. Ubel, P.A. and Asch, D.A. Creating Value In Health By Understanding And Overcoming Resistance To De-Innovation. *Health Aff.* 2015, **34**(2), pp.239-244.
45. Cancer Research UK. Early detection and diagnosis of cancer. A roadmap to the future. [Online]. 2020. [Accessed 06/10/2020]. Available from: https://www.cancerresearchuk.org/sites/default/files/early_detection_diagnosis_of_cancer_roadmap.pdf
46. Brooks, A., Nunes, J.K., Garnett, A., Biellik, R., Leboulleux, D., Birkett, A.J. and Loucq, C. Aligning new interventions with developing country health systems: Target product profiles, presentation, and clinical trial design. *Glob Public Health.* 2012, **7**(9), pp.931-945.
47. Tebbey, P.W. and Rink, C. Target Product Profile: A Renaissance for its Definition and Use. *Int J Med Mark.* 2009, **9**(4), pp.301-307.
48. US Food and Drug Administration. Guidance for industry and review staff Target Product Profile — A strategic development process tool. [Online]. 2007. [Accessed 15/11/2018]. Available from: https://www.fda.gov/media/72566/download
49. Lambert, W.J. Considerations in Developing a Target Product Profile for Parenteral Pharmaceutical Products. *AAPS PharmSciTech.* 2010, **11**(3), pp.1476-1481.
50. Breder, C.D., Du, W. and Tyndall, A. What's the Regulatory Value of a Target Product Profile? *Trends Biotechnol.* 2017, **35**(7), pp.576-579.
51. Tyndall, A., Du, W. and Breder, C.D. The target product profile as a tool for regulatory communication: advantageous but underused. *Nat Rev Drug Discov.* 2017, **16**(3), pp.156-156.
52. International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use. ICH Official website. [Online]. [Accessed 20/07/2022]. Available from: https://www.ich.org/page/mission
53. European Medicines Agency. International Conference on Harmonisation of Technical Requirements for Registration for Pharmaceuticals for Human Use Considerations (ICH) Guideline Q8 (R2) on Pharmaceutical Development. [Online]. 2009. [Accessed 24/03/2022]. Available from: https://www.ema.europa.eu/en/documents/scientific-guideline/international-conference-harmonisation-technical-requirements-registration-pharmaceuticals-human-use_en-11.pdf
54. US Food and Drug Administration. Guidance for industry: Q8 (2) pharmaceutical development. [Online]. 2009. [Accessed 12/02/2019]. Available from: http://academy.gmp-compliance.org/guidemgr/files/9041FNL%5b1%5d.PDF
55. Namjoshi, S., Dabbaghi, M., Roberts, M.S., Grice, J.E. and Mohammed, Y. Quality by Design: Development of the Quality Target Product Profile (QTPP) for Semisolid Topical Products. *Pharmaceutics.* 2020, **12**(3).
56. Cook, J., Cruañes, M.T., Gupta, M., Riley, S. and Crison, J. Quality-by-Design: Are We There Yet? *AAPS PharmSciTech.* 2014, **15**(1), pp.140-148.

57. WHO, FIND and MsF. A multiplex multi-analyte diagnostic platform. [Online]. 2017. [Accessed 08/11/2018]. Available from: https://www.who.int/medicines/TPP_intro_20171122_forDistribution.pdf
58. Lee, B.Y. and Burke, D.S. Constructing target product profiles (TPPs) to help vaccines overcome post-approval obstacles. *Vaccine.* 2010, **28**(16), pp.2806-2809.
59. The malE. R. A. Consultative Group on Vaccines. A Research Agenda for Malaria Eradication: Vaccines. *PLoS Med.* 2011, **8**(1), p.e1000398.
60. WHO. High-priority target product profiles for new tuberculosis diagnostics: report of a consensus meeting. [Online]. 2014. [Accessed 08/11/2018]. Available from: https://apps.who.int/iris/bitstream/handle/10665/135617/WHO_HTM_TB_2014.18_eng.pdf?sequence=1
61. Drummond, M., Sculpher, M.J., Claxton, K., Stoddart, G.L. and Torrance, G.W. *Methods for the economic evaluation of health care programmes.* Fourth ed. Oxford, United Kingdom: Oxford University Press, 2015.
62. National Institute for Health and Care Excellence. Guide to the methods of technology appraisal 2013. [Online]. 2013. [Accessed 05/07/2022]. Available from: https://www.nice.org.uk/process/pmg9/resources/guide-to-the-methods-of-technology-appraisal-2013-pdf-2007975843781
63. Briggs, A., Claxton, K. and Sculpher, M.J. *Decision Modelling for Health Economic Evaluation.* Oxford University Press, 2006.
64. Claxton, K., Martin, S., Soares, M., Rice, N., Spackman, E., Hinde, S., Devlin, N., Smith, P.C. and Sculpher, M. Methods for the estimation of the National Institute for Health and Care Excellence cost-effectiveness threshold. *Health Technol Assess.* 2015, **19**(14), pp.1-503, v-vi.
65. York Health Economics Consortium. Net Monetary Benefit. [Online]. 2016. [Accessed 22/01/2020]. Available from: https://yhec.co.uk/glossary/net-monetary-benefit/
66. Pletcher, M.J. and Pignone, M. Evaluating the clinical utility of a biomarker: a review of methods for estimating health impact. *Circ J.* 2011, (1524-4539 (Electronic)).
67. Annemans, L., Genesté, B. and Jolain, B. Early Modelling for Assessing Health and Economic Outcomes of Drug Therapy. *Value Health.* 2000, **3**(6), pp.427-434.
68. Hartz, S. and John, J. Contribution of economic evaluation to decision making in early phases of product development: A methodological and empirical review. *Int J Technol Assess Health Care.* 2008, **24**(4), pp.465-472.
69. Markiewicz, K., van Til, J.A. and Ijzerman, M.J. MEDICAL DEVICES EARLY ASSESSMENT METHODS: SYSTEMATIC LITERATURE REVIEW. *Int J Technol Assess Health Care.* 2014, **30**(2), pp.137-146.
70. Ijzerman, M.J., Koffijberg, H., Fenwick, E. and Krahn, M. Emerging Use of Early Health Technology Assessment in Medical Product Development: A Scoping Review of the Literature. *PharmacoEconomics.* 2017, **35**(7), pp.727-740.
71. Frempong, S.N., Sutton, A.J., Davenport, C. and Barton, P. Economic evaluation of medical tests at the early phases of development: a systematic review of empirical studies. *Expert Rev Pharmacoecon Outcomes Res.* 2018, **18**(1), pp.13-23.
72. Buisman, L.R., Rutten-van Mölken, M.P.M.H., Postmus, D., Luime, J.J., Uyl-de Groot, C.A. and Redekop, W.K. THE EARLY BIRD CATCHES

THE WORM: EARLY COST-EFFECTIVENESS ANALYSIS OF NEW MEDICAL TESTS. *Int J Technol Assess Health Care.* 2016, **32**(1-2), pp.46-53.
73. Graziadio, S., Winter, A., Lendrem, B.C., Suklan, J., Jones, W.S., Urwin, S.G., O'Leary, R.A., Dickinson, R., Halstead, A., Kurowska, K., Green, K., Sims, A., Simpson, A.J., Power, H.M. and Allen, A.J. How to Ease the Pain of Taking a Diagnostic Point of Care Test to the Market: A Framework for Evidence Development. *Micromachines.* 2020, **11**(3).
74. Price, C.P. and St John, A. Anatomy of a value proposition for laboratory medicine. *Clin Chim Acta.* 2014, **436**, pp.104-111.
75. Abel, L., Dakin, H.A., Roberts, N., Ashdown, H.F., Butler, C.C., Hayward, G., Van den Bruel, A., Turner, P.J. and Yang, Y. Is stratification testing for treatment of chronic obstructive pulmonary disease exacerbations cost-effective in primary care? an early cost-utility analysis. *Int J Technol Assess Health Care.* 2019, **35**(2), pp.116-125.
76. Dowdy, D., O'Brien, M. and Bishai, D. Cost-effectiveness of novel diagnostic tools for the diagnosis of tuberculosis. *Int J Tuberc Lung Dis.* 2008, **12**(9), pp.1021-1029.
77. Frempong, S.N., Sutton, A.J., Davenport, C. and Barton, P. Early Economic Evaluation to Identify the Necessary Test Characteristics of a New Typhoid Test to be Cost Effective in Ghana. *PharmacoEconomics.* 2019, **4**, pp.143-147.
78. Frempong, S.N., Shinkins, B., Howdon, D., Messenger, M., Neal, R.D. and Sagoo, G.S. Early economic evaluation of an intervention to improve uptake of the NHS England Diabetes Prevention Programme. *Expert Rev Pharmacoecon Outcomes Res.* 2021, pp.1-11.
79. Fridhammar, A., Axelsson, U., Persson, U., Bjartell, A. and Borrebaeck, C.A.K. The Value of a New Diagnostic Test for Prostate Cancer: A Cost-Utility Analysis in Early Stage of Development. *PharmacoEconomics - Open.* 2021, **5**(1), pp.77-88.
80. Haug, U., Knudsen, A.B., Lansdorp-Vogelaar, I. and Kuntz, K.M. Development of new non-invasive tests for colorectal cancer screening: The relevance of information on adenoma detection. *Int J Cancer.* 2015, **136**(12), pp.2864-2874.
81. Jenniskens, K., Lagerweij, G.R., Naaktgeboren, C.A., Hooft, L., Moons, K.G.M., Poldervaart, J.M., Koffijberg, H. and Reitsma, J.B. Decision analytic modeling was useful to assess the impact of a prediction model on health outcomes before a randomized trial. *J Clin Epidemiol.* 2019, **115**, pp.106-115.
82. Kluytmans, A., Deinum, J., Jenniskens, K., van Herwaarden Antonius, E., Gloerich, J., van Gool Alain, J., van der Wilt Gert, J. and Grutters Janneke, P.C. Clinical biomarker innovation: when is it worthwhile? *Clin Chem Lab Med.* 2019, **57**(11), p.1712.
83. Lansdorp-Vogelaar, I., van Ballegooijen, M., Zauber, A.G., Boer, R., Wilschut, J. and Habbema, J.D.F. At what costs will screening with CT colonography be competitive? A cost-effectiveness approach. *Int J Cancer.* 2009, **124**(5), pp.1161-1168.
84. Lansdorp-Vogelaar, I., Goede, S.L., Bosch, L.J.W., Melotte, V., Carvalho, B., van Engeland, M., Meijer, G.A., de Koning, H.J. and van Ballegooijen, M. Cost-effectiveness of High-performance Biomarker Tests vs Fecal Immunochemical Test for Noninvasive Colorectal Cancer Screening. *Clin Gastroenterol Hepatol.* 2017, **16**(4), pp.504-512.

85. Sutton, A.J., Lamont, J.V., Evans, R.M., Williamson, K., O'Rourke, D., Duggan, B., Sagoo, G.S., Reid, C.N. and Ruddock, M.W. An early analysis of the cost-effectiveness of a diagnostic classifier for risk stratification of haematuria patients (DCRSHP) compared to flexible cystoscopy in the diagnosis of bladder cancer. *PLOS ONE.* 2018, **13**(8), p.e0202796.
86. Trentham-Dietz, A., Ergun, M.A., Alagoz, O., Stout, N.K., Gangnon, R.E., Hampton, J.M., Dittus, K., James, T.A., Vacek, P.M., Herschorn, S.D., Burnside, E.S., Tosteson, A.N.A., Weaver, D.L. and Sprague, B.L. Comparative effectiveness of incorporating a hypothetical DCIS prognostic marker into breast cancer screening. *Breast Cancer Res Treat.* 2018, **168**(1), pp.229-239.
87. van't Hoog, A.H., Cobelens, F., Vassall, A., van Kampen, S., Dorman, S.E., Alland, D. and Ellner, J. Optimal Triage Test Characteristics to Improve the Cost-Effectiveness of the Xpert MTB/RIF Assay for TB Diagnosis: A Decision Analysis. *PLOS ONE.* 2013, **8**(12), p.e82786.
88. York Health Economics Consortium. Sensitivity Analysis. [Online]. 2016. [Accessed 22 Jan 2020]. Available from: https://yhec.co.uk/glossary/sensitivity-analysis/
89. RID-AMR@Leeds. Rapid infection diagnostics to combat antimicrobial resistance. [Online]. 2021. [Accessed 02/12/2021]. Available from: https://ridamr.leeds.ac.uk/
90. Davies, K., Davis, G., Barbut, F., Eckert, C., Petrosillo, N. and Wilcox, M.H. Variability in testing policies and impact on reported Clostridium difficile infection rates: results from the pilot Longitudinal European Clostridium difficile Infection Diagnosis surveillance study (LuCID). *Eur J Clin Microbiol Infect Dis.* 2016, **35**(12), pp.1949-1956.
91. Cocco, P., Ayaz-Shah, A., Shinkins, B., Messenger, M. and West, R. Methods adopted to develop a Target Product Profile (TPP) in the field of medical tests: a systematic review. [Online]. 2018. [Accessed 26/01/2019]. Available from: https://www.crd.york.ac.uk/prospero/display_record.php?ID=CRD42018115133
92. Godin, K., Stapleton, J., Kirkpatrick, S.I., Hanning, R.M. and Leatherdale, S.T. Applying systematic review search methods to the grey literature: a case study examining guidelines for school-based breakfast programs in Canada. 2015, (2046-4053 (Electronic)), p.4:138.
93. McHugh, M.L. Interrater reliability: the kappa statistic. *Biochem Med.* 2012, **22**(3), pp.276-282.
94. US Department of Health and Human Services, US Food and Drug Administration, Center for Drug Evaluation and Research and Center for Veterinary Medicine. Bioanalytical method validation: guidance for industry. [Online]. 2018. [Accessed 13/02/2019]. Available from: https://www.fda.gov/media/70858/download
95. Association for the Advancement of Medical Instrumentation. Human factors engineering- Design of medical devices. [Online]. 2013. [Accessed 03/02/2019]. Available from: https://my.aami.org/aamiresources/previewfiles/he75_1311_preview.pdf
96. International Standards Organization. Terms and definitions in ISO 14001: 2015 - where did they originate from. [Online]. 2015. [Accessed 13/02/2019]. Available from: https://committee.iso.org/files/live/sites/tc207sc1/files/Terms%20and%20

definitions%20in%20ISO%2014001_2015%20-%20where%20did%20they%20originate%20from.pdf
97. Abuhav, I. *ISO 13485:2016: a complete guide to quality management in the medical device industry. .* 2018.
98. Chua, A., Prat, I., Nuebling, C., Wood, D. and Moussy, F. Update on Zika diagnostic tests and WHO's related activities. *PLoS Negl Trop Dis.* 2017, **11**(2), pp.1-2.
99. Denkinger, C., Dolinger, D., Schito, M., Wells, W., Cobelens, F., Pai, M., Zignol, M., Cirillo, D., Alland, D., Casenghi, M., Gallarda, J., Boehme, C. and Perkins, M. Target product profile of a molecular drug-susceptibility test for use in microscopy centers. *J Infect Dis.* 2015, (211), pp.S39-S49.
100. Denkinger, C., Kik, S., Cirillo, D., Casenghi, M., Shinnick, T., Weyer, K., Gilpin, C., Boehme, C., Schito, M., Kimerling, M. and Pai, M. Defining the needs for next generation assays for tuberculosis. *J Infect Dis.* 2015, **211**, pp.S29-38.
101. DIAMETER Project and PATH. Target Product Profile: point-of-care malaria infection detection test for rapid detection of low-density, subclinical malaria infections. [Online]. 2014. [Accessed]. Available from: https://www.finddx.org/wp-content/uploads/2019/03/DIAMETER-poc-malaria-idt-TPP-2014.pdf
102. Ding, X., Ade, M., Baird, J., Cheng, Q., Cunningham, J., Dhorda, M., Drakeley, C., Felger, I., Gamboa, D., Harbers, M., Herrera, S., Lucchi, N., Mayor, A., Mueller, I., Sattabongkot, J., Ratsimbason, A., Richards, J., Tanner, M. and González, I. Defining the next generation of Plasmodium vivax diagnostic tests for control and elimination: Target product profiles. *PLoS Negl Trop Dis.* 2017, **11**(4), pp.1-15.
103. Dittrich, S., Tadesse, B., Moussy, F., Chua, A., Zorzet, A., Tangden, T., Dolinger, D., Page, A., Crump, J., Valerie, D., Bassat, Q., Lubell, Y., Heinrich, N., Newton, P., Rodwell, T. and Gonzalez, I. Target Product Profile for a diagnostic assay to differentiate between bacterial and non-bacterial infections to guide antimicrobials use in resource-limited settings: an expert consensus. *Am J Trop Med Hyg.* 2017, **95**(5), pp.527-527.
104. Donadeu, M., Fahrion, A., Olliaro, P. and Abela-Ridder, B. Target product profiles for the diagnosis of Taenia solium taeniasis, neurocysticercosis and porcine cysticercosis. *PLoS Negl Trop Dis.* 2017, **11**(9), pp.1-18.
105. Ebels, K., Clerk, C., Crudder, C., McGray, S., Magnuson, K., Tietje, K. and Labarre, P. Incorporating user needs into product development for improved infection detection for malaria elimination programs. *IEEE 2014 Global Humanitarian Technology Conference.* 2014, pp.555-560.
106. FIND and Forum for Collaborative HIV Research. High-priority target product profile for hepatitis C diagnosis in decentralized settings: report of a consensus meeting. [Online]. 2015. [Accessed 08/11/2018]. Available from: https://www.finddx.org/wp-content/uploads/2019/03/HCV-TPP-Report_FIND-2015.pdf
107. FIND. Target Product Profile for tests for recent HIV infection. [Online]. 2017. [Accessed 08/11/2018]. Available from: https://www.finddx.org/wp-content/uploads/2019/03/HIV-Incidence-TPP-FIND-2017.pdf
108. FIND. Target Product Profile: rapid test for diagnosis of malaria and screening for hyman African trypanosomiasis (HAT). [Online]. 2017.

[Accessed 08/11/2018]. Available from: https://www.finddx.org/wp-content/uploads/2019/03/TPP-5-HAT-malaria-combo-test_02FEB17.pdf
109. FIND. Target Product Profile: screening test for human African trypanosomiasis (HAT). [Online]. 2017. [Accessed 08/11/2018]. Available from: https://www.finddx.org/wp-content/uploads/2019/03/TPP-1-HAT-screening-test_02FEB17.pdf
110. Gal, M., Francis, N., Hood, K., Villacian, J., Goossens, H., Watkins, A. and Butler, C. Matching diagnostics development to clinical need: Target Product Profile development for a point of care test for community-acquired lower respiratory tract infection. *PLOS ONE.* 2018, **13**(8), pp.1-13.
111. International Diagnostics Centre. Target Product Profile: Point-of-care HIV viral load test. [Online]. 2013. [Accessed 08/11/2018]. Available from: http://www.idc-dx.org/resources/target-product-profile-point-of-care-hiv-viral-load-test
112. International Diagnostics Centre. Target Product Profile (TPP) – Point-of-Care CD4 Test. [Online]. 2013. [Accessed 08/11/2018]. Available from: http://www.idc-dx.org/resources/target-product-profile-point-of-care-cd4-test
113. International Diagnostics Centre. Target Product Profile (TPP) – Point-of-Care EID Test. [Online]. 2013. [Accessed 08/11/2018]. Available from: http://www.idc-dx.org/resources/target-product-profile-point-of-care-eid-test
114. International Diagnostics Centre. Target Product Profile (TPP) – Combined HIV/Syphilis Test. [Online]. 2014. [Accessed 08/11/2018]. Available from: http://www.idc-dx.org/resources/target-product-profile-combined-hivsyphilis-test
115. Lim, M., Brooker, S., Belizario, V.J., Gay-Andrieu, F., Gilleard, J., Levecke, B., van Lieshout, L., Medley, G., Mekonnen, Z., Mirams, G., Njenga, S., Odiere, M., Rudge, J., Stuyver, L., Vercruysse, J., Vlaminck, J. and Walson, J. Diagnostic tools for soil-transmitted helminths control and elimination programs: a pathway for diagnostic product development. *PLoS Negl Trop Dis.* 2018, **12**(3), pp.1-18.
116. Nsanzabana, C., Ariey, F., Beck, H., Ding, X., Kamau, E., Krishna, S., Legrand, E., Lucchi, N., Miotto, O., Nag, S., Noedl, H., Roper, C., Rosenthal, P.J., Schallig, H., Taylor, S., Volkman, S. and Gonzalez, I. Molecular assays for antimalarial drug resistance surveillance: A target product profile. *PLOS ONE.* 2018, **13**(9), pp.1-18.
117. Pal, S., Jasper, L., Lawrence, K., Walter, M., Gilliland, T., Dauner, A., Palys, T. and Wu, S. Assessing the dengue diagnosis capability gap in the military health system. *Mil Med.* 2016, **181**(8), pp.756-766.
118. PATH. Target Product Profile: HIV self-test Version 4.1: a white paper on the evaluation of current HIV rapid tests and development of core specifications for next-generation HIV tests. [Online]. 2014. [Accessed 08/11/2018]. Available from: https://path.azureedge.net/media/documents/TS_hiv_self_test_tpp.pdf
119. PATH. Diagnostics for neglected tropical diseases: defining the best tools through target product profiles. [Online]. 2015. [Accessed 08/11/2018]. Available from: https://path.azureedge.net/media/documents/DIAG_ntd_prod_profile_rpt.pdf

120. PATH. Target Product Profile: schistosomiasis surveillance diagnostic. [Online]. 2015. [Accessed 08/11/2018]. Available from: https://path.azureedge.net/media/documents/2015.01.15_BMGF_SCH_MDAreduction_NoTech.pdf
121. PATH. Target Product Profile: schistosomiasis surveillance diagnostic - schistosoma genus-specific antibody lateral flow test post-elimination surveillance. [Online]. 2015. [Accessed 08/11/2018]. Available from: https://path.azureedge.net/media/documents/2015.01.15_BMGF_SCH_postMDA_Ab.pdf
122. PATH. Target Product Profile: schistosomiasis surveillance diagnostic lateral flow test CAA antigen. [Online]. 2015. [Accessed 08/11/2018]. Available from: https://path.azureedge.net/media/documents/2015.01.15_BMGF_SCH_MDAreduction_Ag.pdf
123. PATH. Target Product Profile: trachoma surveillance diagnostic - antigen lateral flow test. [Online]. 2015. [Accessed 08/11/2018]. Available from: https://path.azureedge.net/media/documents/2015.01.15_BMGF_Trachoma_MDAreduction_Ag.pdf
124. PATH. Target Product Profile: trachoma surveillance diagnostic - Post-elimination surveillance antibody lateral flow test. [Online]. 2015. [Accessed 08/11/2018]. Available from: https://path.azureedge.net/media/documents/2015.01.15_BMGF_Trachoma_postMDA_Ab.pdf
125. Peck, R., Wellhausen, J., Boettcher, M. and Downing, L. Defining product requirements for field deployable yellow fever tests. [Online]. 2012. [Accessed 08/11/2018]. Available from: http://www.ajtmh.org/docserver/fulltext/14761645/89/5_Suppl_1/ASTMH_13_Abstracts_1001_1250.pdf?expires=1565877435&id=id&accname=guest&checksum=5DBD842D2A8FA98A39884EB2A0713FA3
126. Porras, A., Yadon, Z., Altcheh, J., Britto, C., Chaves, G., Flevaud, L., Martins, O., Ribeiro, I., Schijman, A., Shikanai-Yasuda, M., Sosa-Estani, S., Stobbaerts, E. and Zicker, F. Target Product Profile (TPP) for Chagas disease point-of-care diagnosis and assessment of response to treatment. *PLoS Negl Trop Dis.* 2015, **9**(6), pp.1-8.
127. Reipold, E., Easterbrook, P., Trianni, A., Panneer, N., Krakower, D., Ongarello, S., Roberts, T., Miller, V. and Denkinger, C. Optimising diagnosis of viraemic hepatitis C infection: the development of a target product profile. *BMC Infect Dis.* 2017, **17**, pp.159-175.
128. Solomon, A., Engels, D., Bailey, R., Blake, I., Brooker, S., Chen, J., Chen, J., Churcher, T., Drakeley, C., Edwards, T., Fenwick, A., French, M., Gabrielli, A., Grassly, N., Harding-Esch, E., Holland, M., Koukounari, A., Lammie, P., Leslie, J., Mabey, D., Rhajaoui, M., Secor, W., Stothard, J., Wei, H., Willingham, A., Zhou, X. and Peeling, R. A diagnostics platform for the integrated mapping, monitoring, and surveillance of neglected tropical diseases: Rationale and target product profiles. *PLoS Negl Trop Dis.* 2012, **6**(7), pp.1-5.
129. Toskin, I., Murtagh, M., Peeling, R., Blondeel, K., Cordero, J. and Kiarie, J. Advancing prevention of sexually transmitted infections through point-of-care testing: target product profiles and landscape analysis. *Sex Transm Infect.* 2017, **93**(S4), pp.S69-S80.
130. UNICEF. Pneumonia acute respiratory infection diagnostic aid - Target Product Profile introduction. [Online]. 2014. [Accessed 08/11/2018].

Available from: https://www.unicef.org/supply/files/Pneumonia_Diagnostics_Aid_Device_TPP_Introduction.pdf
131. UNICEF. UNICEF Target Product Profile: real time E. coli detection (version 1.4). . [Online]. 2016. [Accessed 08/11/2018]. Available from: https://www.unicef.org/supply/files/UNICEF_TPP__Real_Time_E_coli_Detection__22_September.pdf
132. UNICEF. UNICEF Target Product Profile: rapid E. coli detection (version 2.0). [Online]. 2017. [Accessed 08/11/2018]. Available from: https://www.unicef.org/supply/files/UNICEF_TPP_Rapid_E._coli_Detection_v2.2_Final_2017.08.04.pdf
133. UNICEF. Zika virus diagnostics: Target Product Profiles & supply update. [Online]. 2017. [Accessed 08/11/2018]. Available from: https://www.unicef.org/supply/files/Zika_Virus_Diagnostics_Target_Product_Profile_and_Supply_Note.pdf
134. Utzinger, J., Becker, S., van Lieshout, L., van Dam, G. and Knopp, S. New diagnostic tools in schistosomiasis. *Clin Microbiol Infect.* 2015, **21**(6), pp.529-542.
135. WHO. Taenia solium taeniasis/cysticercosis diagnostic tool. Report of a stakeholder meeting. [Online]. 2015. [Accessed 08/11/2018]. Available from: https://apps.who.int/iris/bitstream/handle/10665/206543/9789241510516_eng.pdf?sequence=1&isAllowed=y
136. WHO and FIND. Target Product Profile for Zaire ebolavirus rapid, simple test to be used in the control of the Ebola outbreak in West Africa. [Online]. 2014. [Accessed 08/11/2018]. Available from: https://www.who.int/medicines/publications/target-product-profile.pdf?ua=1
137. WHO and FIND. Development of a Target Product Profile (TPP) and a framework for evaluation for a test for predicting progression from tubercolosis infection to active disease. [Online]. 2017. [Accessed 08/11/2018]. Available from: https://apps.who.int/iris/bitstream/handle/10665/259176/WHO-HTM-TB-2017.18-eng.pdf;jsessionid=BD8E426CA555DE9FEAB589FF2A9B2CCC?sequence=1
138. WHO and FIND. Report of a WHO-FIND meeting on diagnostics for Buruli ulcer. [Online]. 2018. [Accessed 08/11/2018]. Available from: https://apps.who.int/iris/bitstream/handle/10665/274607/WHO-CDS-NTD-IDM-2018.09-eng.pdf?sequence=1&isAllowed=y
139. WHO. Specifications for a rapid diagnostic test for meningitis african meningitis belt. [Online]. 2016. [Accessed 08/11/2018]. Available from: https://www.who.int/csr/disease/meningococcal/TPP-Meningitis-RDT-Apr16.pdf?ua=1
140. Slater, H.C., Ross, A., Ouédraogo, A.L., White, L.J., Nguon, C., Walker, P.G.T., Ngor, P., Aguas, R., Silal, S.P., Dondorp, A.M., La Barre, P., Burton, R., Sauerwein, R.W., Drakeley, C., Smith, T.A., Bousema, T. and Ghani, A.C. Assessing the impact of next-generation rapid diagnostic tests on Plasmodium falciparum malaria elimination strategies. *Nature.* 2015, **528**(7580), pp.S94-S101.
141. Sun, A.Y., Pai, M., Salje, H., Satyanarayana, S., Deo, S. and Dowdy, D.W. Modeling the Impact of Alternative Strategies for Rapid Molecular

Diagnosis of Tuberculosis in Southeast Asia. *Am J Epidemiol.* 2013, **178**(12), pp.1740-1749.
142. Batwala, V., Magnussen, P., Hansen, K.S. and Nuwaha, F. Cost-effectiveness of malaria microscopy and rapid diagnostic tests versus presumptive diagnosis: implications for malaria control in Uganda. *Malar J.* 2011, **10**(1), p.372.
143. Assefa, L.M., Crellen, T., Kepha, S., Kihara, J.H., Njenga, S.M., Pullan, R.L. and Brooker, S.J. Diagnostic Accuracy and Cost-Effectiveness of Alternative Methods for Detection of Soil-Transmitted Helminths in a Post-Treatment Setting in Western Kenya. *PLoS Negl Trop Dis.* 2014, **8**(5), p.e2843.
144. Murtagh, M., Blondeel, K., Peeling, R.W., Kiarie, J. and Toskin, I. The relevance of target product profiles for manufacturers, experiences from the World Health Organization initiative for point-of-care testing for sexually transmitted infections. *Arch Public Health.* 2021, **79**(1), p.187.
145. Kaplan, W., Wirtz, V., Mantel-Teeuwisse, A., Stolk, P., Duthey, B. and Laing, R. Priority medicines for Europe and the world- 2013 update [Online]. 2013. [Accessed 10/06//2019]. Available from: https://www.who.int/medicines/areas/priority_medicines/MasterDocJune28_FINAL_Web.pdf?ua=1
146. WHO. An R&D Blueprint for action to prevent epidemics [Online]. 2016. [Accessed 10/06/2019]. Available from: https://www.who.int/blueprint/about/r_d_blueprint_plan_of_action.pdf?ua=1
147. Mundel, T. How do you decide what to invest in? Bill & Melinda Gates Foundation. [Online]. 2018. [Accessed 17/02/2022]. Available from: https://www.gatesfoundation.org/ideas/articles/how-do-you-decide-what-to-invest-in
148. World Health Organisation. Health innovation for impact. [Online]. [Accessed 26/03/2022]. Available from: https://www.who.int/teams/digital-health-and-innovation/health-innovation-for-impact
149. Ruger, J.P. and Yach, D. The Global Role of the World Health Organization. *Glob Health Gov.* (1939-2389 (Print)).
150. Lisk, F. and Šehović, A.B. Rethinking Global Health Governance in a Changing World Order for Achieving Sustainable Development: The Role and Potential of the 'Rising Powers'. *Fudan J Humanit Soc Sci.* 2020, **13**(1), pp.45-65.
151. Tversky, A. and Kahneman, D. Judgment under uncertainty: heuristics and biases. *Science.* 1974, **185**(4157), pp.1124-1131.
152. Pallier, G., Wilkinson, R., Danthiir, V., Kleitman, S., Knezevic, G., Stankov, L. and Roberts, R.D. The Role of Individual Differences in the Accuracy of Confidence Judgments. *J Gen Psychol.* 2002, **129**(3), pp.257-299.
153. Whiting, P., Davenport, C., Jameson, C., Burke, M., Sterne, J., Hyde, C. and Ben-Shlomo, Y. How well do health professionals interpret diagnostic information? A systematic review. *BMJ Open.* 2015, **5**(7), pp.1-8.
154. Linnet, K., Bossuyt, P., Moons, K. and Reitsma, J. Quantifying the accuracy of a diagnostic test or marker. *Clin Chem.* 2012, **58**(9), pp.1292-1301.

155. Breheny, K., Sutton, A.J. and Deeks, J.J. Model-based economic evaluations of diagnostic point of care tests were rarely fit for purpose. *J Clin Epidemiol.* 2019, **109**, pp.1-11.
156. Garfield, S., Polisena, J., Spinner, D., Postulka, A., Lu, C., Tiwana, S., Faulkner, E., Poulios, N., Zah, V. and Longacre, M. Health Technology Assessment for molecular diagnostics: practices, challenges, and recommendations from the Medical Devices and Diagnostics Special Interest Group. *Value Health.* 2016, **19**(5), pp.577-587.
157. World Health Organisation. A Coordinated Global Research Roadmap: 2019 Novel Coronavirus. [Online]. 2020. [Accessed 06/11/2021]. Available from: https://www.who.int/publications/m/item/a-coordinated-global-research-roadmap
158. World Health Organisation. WHO Target Product Profiles for COVID-19 Vaccines. [Online]. 2020. [Accessed 06/11/2021]. Available from: https://www.who.int/publications/m/item/who-target-product-profiles-for-covid-19-vaccines
159. World Health Organisation. Target product profiles for priority diagnostics to support response to the COVID-19 pandemic v.1.0. [Online]. 2020. [Accessed 13/11/2021]. Available from: https://cdn.who.int/media/docs/default-source/blue-print/who-rd-blueprint-diagnostics-tpp-final-v1-0-28-09-jc-ppc-final-cmp92616a80172344e4be0edf315b582021.pdf?sfvrsn=e3747f20_1&download=true
160. FINDDx. SARS-COV-2 DIAGNOSTICS: PERFORMANCE DATA. [Online]. [Accessed 06/11/2021]. Available from: https://www.finddx.org/covid-19/dx-data/
161. Ghaffari, A., Meurant, R. and Ardakani, A. COVID-19 Point-of-Care Diagnostics That Satisfy Global Target Product Profiles. *Diagnostics.* 2021, **11**(1).
162. Medicines & Healthcare products Regulatory Agency. Target Product Profile antibody tests to help determine if people have immunity to SARS-CoV-2 (Version 2.0). [Online]. 2020. [Accessed 24/04/2020]. Available from: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/881162/Target_Product_Profile_antibody_tests_to_help_determine_if_people_have_immunity_to_SARS-CoV-2_Version_2.pdf
163. Department of Health & Social Care. National technical validation process for manufacturers of SARS-CoV-2 (COVID-19) tests. [Online]. 2021. [Accessed 30/10/2021]. Available from: https://www.gov.uk/government/publications/assessment-and-procurement-of-coronavirus-covid-19-tests/coronavirus-covid-19-serology-and-viral-detection-testing-uk-procurement-overview
164. Stevenson, M., Metry, A. and Messenger, M. Modelling of hypothetical SARS-CoV-2 point of care tests on admission to hospital from A&E: rapid cost-effectiveness analysis. 2021, **25**, p.21.
165. National Institute for, H. and Care, E. Exploratory Economic Modelling of SARS-CoV-2 Viral Detection Point of Care Tests and Serology Tests. Final Scope. London: NICE. 2020. [Accessed 19 January 2021]. Available from: www.nice.org.uk/guidance/gid-dg10038/documents/final-scope

166. Miller, M.B., Atrzadeh, F., Burnham, C.A.D., Cavalieri, S., Dunn, J., Jones, S., Mathews, C., McNult, P., Meduri, J., Newhouse, C., Newton, D., Oberholzer, M., Osiecki, J., Pedersen, D., Sweeney, N., Whitfield, N., Campos, J. and Kraft Colleen, S. Clinical Utility of Advanced Microbiology Testing Tools. *J Clin Microbiol.* **57**(9), pp.e00495-00419.
167. Husereau, D., Drummond, M., Petrou, S., Carswell, C., Moher, D., Greenberg, D., Augustovski, F., Briggs, A.H., Mauskopf, J. and Loder, E. Consolidated Health Economic Evaluation Reporting Standards (CHEERS) -- Explanation and Elaboration: A Report of the ISPOR Health Economic Evaluation Publication Guidelines Good Reporting Practices Task Force. *Value in Health.* 2013, **16**(2), pp.231-250.
168. European Centre For Disease Prevention and Control. Point prevalence survey of healthcare-associated infections and antimicrobial use in European acute care hospitals 2011-2012. [Online]. 2013. [Accessed 27/05/2021]. Available from: https://www.ecdc.europa.eu/sites/portal/files/media/en/publications/Publications/healthcare-associated-infections-antimicrobial-use-PPS.pdf
169. Finn, E., Andersson, F.L. and Madin-Warburton, M. Burden of Clostridioides difficile infection (CDI) - a systematic review of the epidemiology of primary and recurrent CDI. *BMC Infect Dis.* 2021, **21**(1), p.456.
170. Rodriguez, C., Van Broeck, J., Taminiau, B., Delmée, M. and Daube, G. Clostridium difficile infection: Early history, diagnosis and molecular strain typing methods. *Microb Pathog.* 2016, **97**, pp.59-78.
171. Eze, P., Balsells, E., Kyaw, M.H. and Nair, H. Risk factors for Clostridium difficile infections - an overview of the evidence base and challenges in data synthesis. *J Glob Health.* 2017, (2047-2986 (Electronic)).
172. Pollock, N.R. Ultrasensitive Detection and Quantification of Toxins for Optimized Diagnosis of Clostridium difficile Infection. *J Clin Microbiol.* 2016, (1098-660X (Electronic)).
173. Lewis, B.B. and Pamer, E.G. Microbiota-Based Therapies for Clostridium difficile and Antibiotic-Resistant Enteric Infections. *Annu Rev Microbiol.* 2017, **71**(1), pp.157-178.
174. Burke, K.E. and Lamont, J.T. Clostridium difficile infection: a worldwide disease. *Gut Liver.* 2014, (2005-1212 (Electronic)).
175. Leffler, D.A. and Lamont, J.T. Clostridium difficile infection. *N Engl J Med.* 2015, (1533-4406 (Electronic)).
176. Bloomfield, L.E. and Riley, T.V. Epidemiology and Risk Factors for Community-Associated Clostridium difficile Infection: A Narrative Review. *Infect Dis Ther.* 2016, (2193-8229 (Print)).
177. Larson, H.E., Price, A.B., Honour, P. and Borriello, S.P. CLOSTRIDIUM DIFFICILE AND THE AETIOLOGY OF PSEUDOMEMBRANOUS COLITIS. *Lancet.* 1978, **311**(8073), pp.1063-1066.
178. Department of Health. Clostridium difficile infection: how to deal with the problem. [Online]. 2008. [Accessed 01/03/2022]. Available from: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/340851/Clostridium_difficile_infection_how_to_deal_with_the_problem.pdf
179. Caroff, D.A., Edelstein, P.H., Hamilton, K., Pegues, D.A. and Gilligan, P.H. The Bristol Stool Scale and Its Relationship to Clostridium difficile Infection. *J Clin Microbiol.* 2014, **52**(9), pp.3437-3439.

180. National Institute for Health and Care Excellence. Clostridioides difficile infection: antimicrobial prescribing. [Online]. 2021. [Accessed 01/03/2022]. Available from: https://www.nice.org.uk/guidance/ng199/resources/clostridioides-difficile-infection-antimicrobial-prescribing-pdf-66142090546117
181. Public Health England. Updated guidance on the management and treatment of Clostridium difficile infection. [Online]. 2013. [Accessed 20/01/20]. Available from: https://www.gov.uk/government/publications/clostridium-difficile-infection-guidance-on-management-and-treatment
182. Bauer, M.P., Notermans, D.W., van Benthem, B.H.B., Brazier, J.S., Wilcox, M.H., Rupnik, M., Monnet, D.L., van Dissel, J.T. and Kuijper, E.J. *Clostridium difficile* infection in Europe: a hospital-based survey. *Lancet.* 2011, **377**(9759), pp.63-73.
183. Eyre, D.W., Walker, A.S., Wyllie, D., Dingle, K.E., Griffiths, D., Finney, J., O'Connor, L., Vaughan, A., Crook, D.W., Wilcox, M.H., Peto, T.E.A. and for the Infections in Oxfordshire Research, D. Predictors of First Recurrence of Clostridium difficile Infection: Implications for Initial Management. *Clin Infect Dis.* 2012, **55**(suppl_2), pp.S77-S87.
184. Cornely, O.A., Miller, M.A., Louie, T.J., Crook, D.W. and Gorbach, S.L. Treatment of First Recurrence of Clostridium difficile Infection: Fidaxomicin Versus Vancomycin. *Clin Infect Dis.* 2012, **55**(suppl_2), pp.S154-S161.
185. Blondeau, J.M. What have we learned about antimicrobial use and the risks for Clostridium difficile-associated diarrhoea? *J Antimicrob Chemother.* 2009, (1460-2091 (Electronic)).
186. Donskey, C.J. Preventing Transmission of Clostridium difficile: Is the Answer Blowing in the Wind? *Clin Infect Dis.* 2010, **50**(11), pp.1458-1461.
187. Hardy, K.J., Gossain, S., Thomlinson, D., Pillay, D.G. and Hawkey, P.M. Reducing Clostridium difficile through early identification of clusters and the use of a standardised set of interventions. *J Hosp Infect.* 2010, **75**(4), pp.277-281.
188. Gerding, D.N., Muto, C.A. and Owens, R.C., Jr. Measures to Control and Prevent Clostridium difficile Infection. *Clin Infect Dis.* 2008, **46**(Supplement_1), pp.S43-S49.
189. Pépin, J., Valiquette, L. and Cossette, B. Mortality attributable to nosocomial Clostridium difficile–associated disease during an epidemic caused by a hypervirulent strain in Quebec. *Canad Med Assoc J.* 2005, **173**(9), p.1037.
190. Micek, S.T., Schramm, G., Morrow, L., Frazee, E., Personett, H., Doherty, J.A., Hampton, N., Hoban, A., Lieu, A., McKenzie, M., Dubberke, E.R. and Kollef, M.H. Clostridium difficile Infection: A Multicenter Study of Epidemiology and Outcomes in Mechanically Ventilated Patients. *Crit Care Med.* 2013, **41**(8).
191. Bouza, E. Consequences of *Clostridium difficile* infection: understanding the healthcare burden. *Clin Microbiol Infect.* 2012, **18**, pp.5-12.
192. Olsen, M.A., Stwalley, D., Demont, C. and Dubberke, E.R. Clostridium difficile infection increases acute and chronic morbidity and mortality. *Infect Control Hosp Epidemiol.* 2019, **40**(1), pp.65-71.

193. Heimann, S.M., Cruz Aguilar, M.R., Mellinghof, S. and Vehreschild, M.J.G.T. Economic burden and cost-effective management of Clostridium difficile infections. *Med Mal Infect.* 2018, **48**(1), pp.23-29.
194. Tresman, R. and Goldenberg, S.D. Healthcare resource use and attributable cost of Clostridium difficile infection: a micro-costing analysis comparing first and recurrent episodes. *J Antimicrob Chemother.* 2018, **73**(10), pp.2851-2855.
195. Wilcox, M.H., Ahir, H., Coia, J.E., Dodgson, A., Hopkins, S., Llewelyn, M.J., Settle, C., McLain-Smith, S. and Marcella, S.W. Impact of recurrent Clostridium difficile infection: hospitalization and patient quality of life. *J Antimicrob Chemother.* 2017, **72**(9), pp.2647-2656.
196. Liubakka, A. and Vaughn, B.P. Clostridium difficile Infection and Fecal Microbiota Transplant. *AACN Adv Crit Care.* 2016, **27**(3), pp.324-337.
197. Polage, C., Gyorke, C.E., Kennedy, M.A., Leslie, J.L., Chin, D.L., Wang, S., Nguyen, H.H., Huang, B., Tang, Y., Lee, L.W., Kim, K., Taylor, S., Romano, P.S., Panacek, E.A., Goodell, P.B., Solnick, J.V. and Cohen, S.H. Overdiagnosis of Clostridium difficile Infection in the Molecular Test Era. *JAMA Intern Med.* 2015, **175**(11), pp.1792-1801.
198. Chen, S., Gu, H., Sun, C., Wang, H. and Wang, J. Rapid detection of Clostridium difficile toxins and laboratory diagnosis of Clostridium difficile infections. *Infection.* 2017, **45**(3), pp.255-262.
199. Kufelnicka, A.M. and Kirn, T.J. Effective utilization of evolving methods for the laboratory diagnosis of Clostridium difficile infection. *Clin Infect Dis.* 2011, **52**(12), pp.1451-1457.
200. Gateau, C., Couturier, J., Coia, J. and Barbut, F. How to: diagnose infection caused by Clostridium difficile. *Clin Microbiol Infect.* 2018, **24**(5), pp.463-468.
201. Pollock, N.R. Ultrasensitive detection and quantification of toxins for optimized diagnosis of Clostridium difficile infection. *Journal of Clinical Microbiology.* 2016, **54**(2), pp.259-264.
202. Dubberke, E.R. and Burnham, C.D. Diagnosis of Clostridium difficile Infection: Treat the Patient, Not the Test. *JAMA Intern Med.* 2015, **175**(11), pp.1801-1802.
203. Cohen, S., Gerding, D., Johnson, S., Kelly, C., Loo, V., McDonald, L., Clifford, Pepin, J. and Wilcox, M. Clinical Practice Guidelines for *Clostridium difficile* Infection in Adults: 2010 Update by the Society for Healthcare Epidemiology of America (SHEA) and the Infectious Diseases Society of America (IDSA). *Infect Control Hosp Epidemiol.* 2010, **31**(5), pp.431-455.
204. Doll, M., Fleming, M., Stevens, M.P. and Bearman, G. Clostridioides difficile–Associated Diarrhea: Infection Prevention Unknowns and Evolving Risk Reduction Strategies. *Curr Infect Dis Rep.* 2019, **21**(1), p.1.
205. Riggs, M.M., Sethi, A.K., Zabarsky, T.F., Eckstein, E.C., Jump, R.L.P. and Donskey, C.J. Asymptomatic Carriers Are a Potential Source for Transmission of Epidemic and Nonepidemic Clostridium difficile Strains among Long-Term Care Facility Residents. *Clin Infect Dis.* 2007, **45**(8), pp.992-998.
206. Wilcox, M.H. Overcoming barriers to effective recognition and diagnosis of Clostridium difficile infection. *Clin Microbiol Infect.* 2012, **18**(SUPPL.6), pp.13-20.

207. Humphries, R.M. Laboratory Tests for the Diagnosis of Clostridium difficile Infections. *Clin Microbiol Newsl.* 2012, **34**(19), pp.151-157.
208. Kociolek, L.K. Strategies for optimizing the diagnostic predictive value of clostridium difficile molecular diagnostics. *J Clin Microbiol.* 2017, **55**(5), pp.1244-1248.
209. Carter, G.P., Chakravorty, A., Pham Nguyen, T.A., Mileto, S., Schreiber, F., Li, L., Howarth, P., Clare, S., Cunningham, B., Sambol, S.P., Cheknis, A., Figueroa, I., Johnson, S., Gerding, D.N., Rood, J.I., Dougan, G., Lawley, T.D. and Lyras, D. Defining the Roles of TcdA and TcdB in Localized Gastrointestinal Disease, Systemic Organ Damage, and the Host Response during *Clostridium difficile* Infections. *mBio.* 2015, **6**(3), pp.e00551-00515.
210. Kuehne, S.A., Cartman, S.T., Heap, J.T., Kelly, M.L., Cockayne, A. and Minton, N.P. The role of toxin A and toxin B in Clostridium difficile infection. *Nature.* 2010, **467**(7316), pp.711-713.
211. Lyras, D., O'Connor, J.R., Howarth, P.M., Sambol, S.P., Carter, G.P., Phumoonna, T., Poon, R., Adams, V., Vedantam, G., Johnson, S., Gerding, D.N. and Rood, J.I. Toxin B is essential for virulence of Clostridium difficile. *Nature.* 2009, **458**(7242), pp.1176-1179.
212. Gilligan, P.H. Contemporary approaches for the laboratory diagnosis of Clostridium difficile infections. *Semin Colon Rectal Surg.* 2014, **25**(3), pp.137-142.
213. Wilcox, M.H., Planche, T. and Fang, F.C. What is the current role of algorithmic approaches for diagnosis of Clostridium difficile infection? *J Clin Microbiol.* 2010, **48**(12), pp.4347-4353.
214. Gerding, D.N., File, T.M. and McDonald, L.C. Diagnosis and treatment of clostridium difficile infection. *Infect Dis Clin Pract.* 2016, **24**(1), pp.3-10.
215. Tenover, F.C., Baron, E.J., Peterson, L.R. and Persing, D.H. Laboratory diagnosis of Clostridium difficile infection can molecular amplification methods move us out of uncertainty? *J Mol Diagn.* 2011, **13**(6), pp.573-582.
216. Martinez-Melendez, A., Camacho-Ortiz, A., Morfin-Otero, R., Maldonado-Garza, H.J., Viiiarreai-Trevino, L. and Garza-Gonzaiez, E. Current knowledge on the laboratory diagnosis of Clostridium difficile infection. *World J Gastroenterol.* 2017, **23**(9), pp.1552-1567.
217. McFarland, L.V. Renewed interest in a difficult disease: Clostridium difficile infections--epidemiology and current treatment strategies. *Curr Opin Gastroenterol.* 2009, **25**(1), pp.24-35.
218. Planche, T.D., Davies, K.A., Coen, P.G., Finney, J.M., Monahan, I.M., Morris, K.A., O'Connor, L., Oakley, S.J., Pope, C.F., Wren, M.W., Shetty, N.P., Crook, D.W. and Wilcox, M.H. Differences in outcome according to *Clostridium difficile* testing method: a prospective multicentre diagnostic validation study of *C difficile* infection. *The Lancet Infectious Diseases.* 2013, **13**(11), pp.936-945.
219. Planche, T. and Wilcox, M. Reference assays for Clostridium difficile infection: one or two gold standards? *J Clin Pathol.* 2011, **64**(1), pp.1-5.
220. Le Monnier, A., Zahar, J.R. and Barbut, F. Update on clostridium difficile infections. *Med Mal Infect.* 2014, **44**(8), pp.354-365.
221. Goldenberg, S.D. and French, G.L. Diagnostic testing for Clostridium difficile: a comprehensive survey of laboratories in England. *J Hosp Infect.* 2011, **79**(1), pp.4-7.

222. Crobach, M.J.T., Planche, T., Eckert, C., Barbut, F., Terveer, E.M., Dekkers, O.M., Wilcox, M.H. and Kuijper, E.J. European Society of Clinical Microbiology and Infectious Diseases: update of the diagnostic guidance document for Clostridium difficile infection. *Clin Microbiol Infect.* 2016, **22 Suppl 4**, pp.S63-81.
223. Department of Health. Updated guidance on the diagnosis and reporting of Clostridium Difficile. [Online]. 2012. [Accessed 18/10/21]. Available from: https://www.gov.uk/government/publications/updated-guidance-on-the-diagnosis-and-reporting-of-clostridium-difficile
224. Del Prete, R., Ronga, L., Addati, G., Magrone, R., Abbasciano, A., Decimo, M. and Miragliotta, G. Clostridium difficile. A review on an emerging infection. *Clin Ter.* 2019, **170**(1), pp.E41-E47.
225. Berg, A.M., Kelly, C.P. and Farraye, F.A. Clostridium difficile infection in the inflammatory bowel disease patient. *Inflamm Bowel Dis.* 2013, **19**(1), pp.194-204.
226. Burnham, C.A.D. and Carroll, K.C. Diagnosis of Clostridium difficile Infection: an Ongoing Conundrum for Clinicians and for Clinical Laboratories. *Clin Microbiol Rev.* 2013, **26**(3), pp.604-630.
227. Cojocariu, C., Stanciu, C., Stoica, O., Singeap, A.M., Sfarti, C., Girleanu, I. and Trifan, A. Clostridium difficile infection and inflammatory bowel disease. *Turk J Gastroenterol.* 2014, **25**(6), pp.603-610.
228. Peng, Z., Ling, L., Stratton, C.W., Li, C., Polage, C.R., Wu, B. and Tang, Y. Advances in the diagnosis and treatment of Clostridium difficile infections. *Emerg Microbes Infect.* 2018, **7**(2), p.15.
229. Czepiel, J., Drozdz, M., Pituch, H., Kuijper, E.J., Perucki, W., Mielimonka, A., Goldman, S., Wultanska, D., Garlicki, A. and Biesiada, G. Clostridium difficile infection: review. *Eur J Clin Microbiol Infect Dis.* 2019.
230. Shetty, N., Wren, M.W.D. and Coen, P.G. The role of glutamate dehydrogenase for the detection of *Clostridium difficile* in faecal samples: a meta-analysis. *J Hosp Infect.* 2011, **77**(1), pp.1-6.
231. Eastwood, K., Else, P., Charlett, A. and Wilcox, M. Comparison of Nine Commercially Available *Clostridium difficile* Toxin Detection Assays, a Real-Time PCR Assay for *C. difficile tcdB* and a Glutamate Dehydrogenase Detection Assay to Cytotoxin Testing and Cytotoxigenic Culture Methods. *J Clin Microbiol.* 2009, **47**(10), p.3211.
232. Usacheva, E.A., Jin, J.P. and Peterson, L.R. Host response to Clostridium difficile infection: Diagnostics and detection. *J Glob Antimicrob Resist.* 2016, **7**, pp.93-101.
233. Barbut, F., Lalande, V. and Eckert, C. New molecular methods for the diagnosis of Clostridium difficile infections. *Drugs Today.* 2012, **48**(10), pp.673-684.
234. Taylor, K.N., McHale, M.T., Saenz, C.C. and Plaxe, S.C. Diagnosis and treatment of Clostridium difficile (C. diff) colitis: Review of the literature and a perspective in gynecologic oncology. *Gynecol Oncol.* 2017, **144**(2), pp.428-437.
235. Bartsch, S.M., Umscheid, C.A., Nachamkin, I., Hamilton, K. and Lee, B.Y. Comparing the economic and health benefits of different approaches to diagnosing Clostridium difficile infection. *Clin Microbiol Infect.* 2015, **21**(1), pp.77.e71-79.
236. Goldenberg, S.D., Bisnauthsing, K.N., Patel, A., Postulka, A., Wyncoll, D., Schiff, R. and French, G.L. Point-of-Care Testing for Clostridium

Difficile Infection: A Real-World Feasibility Study of a Rapid Molecular Test in Two Hospital Settings. *Infect Dis Ther.* 2014, **3**(2), pp.295-306.
237. Dupont, H.L. Diagnosis and management of Clostridium difficile infection. *Clin Gastroenterol Hepatol.* 2013, **11**(10), pp.1216-e1273.
238. Eysenbach, G. Improving the Quality of Web Surveys: The Checklist for Reporting Results of Internet E-Surveys (CHERRIES). *J Med Internet Res.* 2004, **6**(3), p.e34.
239. Bethlehem, J. The Routing Structure of Questionnaires. *Int J Mark Res.* 2000, **42**(1), pp.1-11.
240. Thomas, S.J. Using Web and Paper Questionnaires for Data-Based Decision Making. [Online]. Thousand Oaks, California: SAGE Publications, Inc., 2004. Available from: https://methods.sagepub.com/book/using-web-and-paper-questionnaires-for-data-based-decision-making
241. Van Teijlingen, E. and Hundley, V. The Importance of Pilot Studies. *Nurs Stand.* 2002, **16**, pp.33-36.
242. Ruel, E., Wagner III, W.E. and Gillespie, B.J. The Practice of Survey Research: Theory and Applications. [Online]. Thousand Oaks, California: SAGE Publications, Inc, 2016. Available from: https://methods.sagepub.com/book/the-practice-of-survey-research
243. NIHR Leeds In Vitro Diagnostics Co-operative. [Online]. 2021. [Accessed 02/12/2021]. Available from: https://leedsmic.nihr.ac.uk/
244. Boekhoud, I.M., Sidorov, I., Nooij, S., Harmanus, C., Bos-Sanders, I.M.J.G., Viprey, V., Spittal, W., Clark, E., Davies, K., Freeman, J., Kuijper, E.J., Smits, W.K. and Consortium, C.-C. Haem is crucial for medium-dependent metronidazole resistance in clinical isolates of Clostridioides difficile. *J Antimicrob Chemother.* 2021.
245. Bhattacherjee, A. *Social science research : principles, methods, and practices.* 2012.
246. Kirchherr, J. and Charles, K. Enhancing the sample diversity of snowball samples: Recommendations from a research project on anti-dam movements in Southeast Asia. *PLOS ONE.* 2018, **13**(8), p.e0201710.
247. Lippi, G., Simundic, A., on behalf of the European Federation for Clinical, C. and Laboratory Medicine Working Group for Preanalytical, P. The EFLM strategy for harmonization of the preanalytical phase. *Clin Chem Lab Med.* 2018, **56**(10), pp.1660-1666.
248. Sikaris, K. Performance criteria of the post-analytical phase. *Clin Chem Lab Med.* 2015, **53**(6), pp.949-958.
249. Aguado, J.M., Anttila, V.J., Galperine, T., Goldenberg, S.D., Gwynn, S., Jenkins, D., Norén, T., Petrosillo, N., Seifert, H., Stallmach, A., Warren, T. and Wenisch, C. Highlighting clinical needs in *Clostridium difficile* infection: the views of European healthcare professionals at the front line. *J Hosp Infect.* 2015, **90**(2), pp.117-125.
250. Davies, K., Davis, G., Barbut, F., Eckert, C., Petrosillo, N. and Wilcox, M. Longitudinal European Clostridium difficile Infection Diagnosis Surveillance Study (LuCID) Shows Effects of Place, Patient Age and Testing Method on CDI Reporting. *Open Forum Infect Dis.* 2015, **2**(suppl_1).
251. Davies, K.A., Longshaw, C.M., Davis, G.L., Bouza, E., Barbut, F., Barna, Z., Delmée, M., Fitzpatrick, F., Ivanova, K., Kuijper, E., Macovei, I.S., Mentula, S., Mastrantonio, P., von Müller, L., Oleastro, M., Petinaki, E., Pituch, H., Norén, T., Nováková, E., Nyč, O., Rupnik, M., Schmid, D. and

Wilcox, M.H. Underdiagnosis of *Clostridium difficile* across Europe: the European, multicentre, prospective, biannual, point-prevalence study of *Clostridium difficile* infection in hospitalised patients with diarrhoea (EUCLID). *Lancet Infect Dis.* 2014, **14**(12), pp.1208-1219.
252. Alcalá, L., Martin, A., Marin, M., Sánchez-Somolinos, M., Catalán, P., Peláez, T. and Bouza, E. The undiagnosed cases of *Clostridium difficile* infection in a whole nation: where is the problem? *Clin Microbiol Infect.* 2012, **18**(7), pp.E204-E213.
253. Viprey, V.F., Davis, G.L., Benson, A.D., Ewin, D., Spittal, W., Vernon, J., Rupnik, M., Banz, A., Allantaz, F., Cleuziat, P., Coordinators, C.-C.N., Wilcox, M.H., Davies, K. and COMBACTE-CDI Consortium Group. Key Differences in Diagnosis and Patient Populations between Community and In-Patient Clostridioides Difficile Infections (CDI): Results from Combatting Bacterial Resistance in Europe CDI (COMBACTE-CDI). *Lancet.* 2021.
254. Grein, J.D., Ochner, M., Hoang, H., Jin, A., Morgan, M.A. and Murthy, A.R. Comparison of testing approaches for *Clostridium difficile* infection at a large community hospital. *Clin Microbiol Infect.* 2014, **20**(1), pp.65-69.
255. Vogelzang, E.H., Lankelma, J.M., van Mansfeld, R., van Prehn, J. and van Houdt, R. Implementing a Clostridium difficile testing algorithm and its effect on isolation duration and treatment initiation: a pre- and post-implementation study. *Eur J Clin Microbiol Infect Dis.* 2020, **39**(6), pp.1071-1076.
256. Ferson, K., Montgomery, J., Moore, R.E., Millar, B.C., Leggett, P., Coulter, W.A., Goldsmith, C.E. and Moore, J.E. Reliability of self-reporting of antibiotic consumption in the community – Index of Reliability. *J Clin Pharm Ther.* 2014, **39**(5), pp.468-470.
257. Demoré, B., Le Govic, D., Thilly, N., Boivin, J.M. and Pulcini, C. Reliability of self-reported recent antibiotic use among the general population: a cross-sectional study. *Clin Microbiol Infect.* 2017, **23**(7), pp.486.e487-486.e412.
258. Lippi, G., Banfi, G., Church, S., Cornes, M., De Carli, G., Grankvist, K., KRistensen, G., Ibarz, M., Panteghini, M., Plebani, M., Nybo, M., Smellie, S., Zaninotto, M., Simundic, A. and on behalf of the European Federation for Clinical Chemistry Laboratory Medicine Working Group for Preanalytical Phase. Preanalytical quality improvement. In pursuit of harmony, on behalf of European Federation for Clinical Chemistry and Laboratory Medicine (EFLM) Working group for Preanalytical Phase (WG-PRE). *Clin Chem Lab Med.* 2015, **53**(3), pp.357-370.
259. Plebani, M. Exploring the iceberg of errors in laboratory medicine. *Clin Chim Acta.* 2009, **404**(1), pp.16-23.
260. Barbut, F., Surgers, L., Eckert, C., Visseaux, B., Cuingnet, M., Mesquita, C., Pradier, N., Thiriez, A., Ait-Ammar, N., Aifaoui, A., Grandsire, E. and Lalande, V. Does a rapid diagnosis of *Clostridium difficile* infection impact on quality of patient management? *Clin Microbiol Infect.* 2014, **20**(2), pp.136-144.
261. Amjad, M. An Overview of the Molecular Methods in the Diagnosis of Gastrointestinal Infectious Diseases. *Int J Microbiol.* 2020, **2020**, p.8135724.
262. Greenacre, Z.A. The Importance of Selection Bias in Internet Surveys. *Open J Stat.* 2016, **06**(03), pp.397-404.

263. Davies, K., Viprey, V., Ewin, D., Spittal, W., Vernon, J., Tkalec, V., Fawley, W., Benson, A., Davis, G., Rupnik, M. and Wilcox, M. Burden of Clostridium difficile Infection (CDI) Across Whole Healthcare Economies and European Borders; COMBACTE-CDI Results. *Infect Control Hosp Epidemiol.* 2020, **41**(S1), pp.s24-s25.
264. Bartsch, S.M., Curry, S.R., Harrison, L.H. and Lee, B.Y. The potential economic value of screening hospital admissions for Clostridium difficile. *Eur J Clin Microbiol Infect Dis.* 2012, **31**(11), pp.3163-3171.
265. Freeman, K., Mistry, H., Tsertsvadze, A., Royle, P., McCarthy, N., Taylor-Phillips, S., Manuel, R. and Mason, J. Multiplex tests to identify gastrointestinal bacteria, viruses and parasites in people with suspected infectious gastroenteritis: a systematic review and economic analysis. *Health Technol Assess.* 2017, **21**(23), pp.1-188.
266. Jones, W.S., Rice, S., Power, H.M., Maniatopoulos, G., Suklan, J., Beyer, F., Wilcox, M.H., Permain, M., Simpson, A.J., Price, D.A. and Allen, A.J. Cost Consequences for the NHS of Using a Two-Step Testing Method for the Detection of Clostridium difficile with a Point of Care, Polymerase Chain Reaction Test as the First Step. *Diagnostics.* 2020, **10**(10).
267. Schroeder, L.F., Robilotti, E., Peterson, L.R., Banaei, N. and Dowdy, D.W. Economic evaluation of laboratory testing strategies for hospital-associated Clostridium difficile infection. *J Clin Microbiol.* 2014, **52**(2), pp.489-496.
268. Xuan, S., Zangwill, K.M., Ni, W., Ma, J. and Hay, J.W. Cost-Effectiveness Analysis of Four Common Diagnostic Methods for Clostridioides difficile Infection. *J Gen Intern Med.* 2020, **35**(4), pp.1102-1110.
269. Saab, S., Alper, T., Sernas, E., Pruthi, P., Alper, M.A. and Sundaram, V. Hospitalized Patients with Cirrhosis Should Be Screened for Clostridium difficile Colitis. *Dig Dis Sci.* 2015, **60**(10), pp.3124-3129.
270. Schechner, V., Carmeli, Y. and Leshno, M. A mathematical model of Clostridium difficile transmission in medical wards and a cost-effectiveness analysis comparing different strategies for laboratory diagnosis and patient isolation. *PLOS ONE.* 2017, **12**(2), p.e0171327.
271. Barton, P., Bryan, S. and Robinson, S. Modelling in the economic evaluation of health care: selecting the appropriate approach. *J Health Serv Res Policy.* 2004, **9 (2)**(1355-8196 (Print)), pp.110-118.
272. Roberts, M., Russell, L.B., Paltiel, A.D., Chambers, M., McEwan, P. and Krahn, M. Conceptualizing a Model: A Report of the ISPOR-SMDM Modeling Good Research Practices Task Force–2. *Med Decis Making.* 2012, **32**(5), pp.678-689.
273. Brennan, A., Chick, S.E. and Davies, R. A taxonomy of model structures for economic evaluation of health technologies. *Health Econ.* 2006, **15**(12), pp.1295-1310.
274. Karnon, J. and Brown, J. Selecting a decision model for economic evaluation: a case study and review. *Health Care Manag Sci.* 1998, **1**(2), pp.133-140.
275. Stahl, J.E. Modelling Methods for Pharmacoeconomics and Health Technology Assessment. *PharmacoEconomics.* 2008, **26**(2), pp.131-148.
276. Cooper, K., Brailsford, S.C. and Davies, R. Choice of modelling technique for evaluating health care interventions. *J Oper Res Soc.* 2007, **58**(2), pp.168-176.

277. Squires, H., Chilcott, J., Akehurst, R., Burr, J. and Kelly, M.P. A Framework for Developing the Structure of Public Health Economic Models. *Value in Health.* 2016, **19**(5), pp.588-601.
278. Jin, H., Robinson, S., Shang, W., Achilla, E., Aceituno, D. and Byford, S. Overview and Use of Tools for Selecting Modelling Techniques in Health Economic Studies. *PharmacoEconomics.* 2021, **39**(7), pp.757-770.
279. Nguyen, L., Megiddo, I. and Howick, S. Simulation models for transmission of health care-associated infection: A systematic review. *Am J Infect Control.* 2020, **48**(7), pp.810-821.
280. Briggs, A.H., Weinstein, M.C., Fenwick, E.A.L., Karnon, J., Sculpher, M.J. and Paltiel, A.D. Model Parameter Estimation and Uncertainty: A Report of the ISPOR-SMDM Modeling Good Research Practices Task Force-6. *Value in Health.* 2012, **15**(6), pp.835-842.
281. Leeds Teaching Hospitals NHS Trust. Clostridium Difficile - Prevention of Transmission - in Adults and Children > 2 years with Clostridium Difficile Infection ( CDI ). [Online]. 2017. [Accessed 26/10/2020]. Available from: http://www.lhp.leedsth.nhs.uk/detail.aspx?ID=677
282. Leeds Teaching Hospitals NHS Trust. Isolation - Infection Prevention & Control. [Online]. 2018. [Accessed 26/10/2020]. Available from: http://www.lhp.leedsth.nhs.uk/detail.aspx?ID=678
283. Leeds Teaching Hospitals NHS Trust. Clostridium Difficile Infection ( CDI ) in Adults ( 16 years of age ). [Online]. 2008. [Accessed 24/10/2020]. Available from: http://www.lhp.leedsth.nhs.uk/detail.aspx?ID=1254
284. Leeds Teaching Hospitals NHS Trust. Infections ( Alert Organisms And Conditions ) That Require Source Isolation. [Online]. 2009. [Accessed 26/10/2020]. Available from: http://www.lhp.leedsth.nhs.uk/detail.aspx?id=1671
285. Monks, T., Currie, C.S.M., Onggo, B.S., Robinson, S., Kunc, M. and Taylor, S.J.E. Strengthening the reporting of empirical simulation studies: Introducing the STRESS guidelines. *J Simul.* 2019, **13**(1), pp.55-67.
286. Husereau, D., Drummond, M., Augustovski, F., de Bekker-Grob, E., Briggs, A.H., Carswell, C., Caulley, L., Chaiyakunapruk, N., Greenberg, D., Loder, E., Mauskopf, J., Mullins, C.D., Petrou, S., Pwu, R. and Staniszewska, S. Consolidated Health Economic Evaluation Reporting Standards (CHEERS) 2022 Explanation and Elaboration: A Report of the ISPOR CHEERS II Good Practices Task Force. *Value in Health.* 2022, **25**(1), pp.10-31.
287. Maghdoori, S. and Moghadas, S.M. Assessing the effect of patient screening and isolation on curtailing Clostridium difficile infection in hospital settings. *BMC Infect Dis.* 2017, **17**(1), p.384.
288. Delignette-Muller, M.L. and Dutang, C. fitdistrplus: An R Package for Fitting Distributions. *J Stat Softw.* 2015, **64**(4), pp.1 - 34.
289. Wingen-Heimann, S., Lurienne, L., Davies, K., Benson, A., Davis, G., Viprey, V., Wilcox, M.H., Bonten, M.J.C., Cornely, O.A. and Vehreschild, J.J. Healthcare resource utilisation for treatment of Clostridioides difficile infection across 12 European countries: health economic results of COMBACTE-CDI. In: *ECCMID*. 2020.
290. SIMUL8 Corporation. Probability Profile. [Online]. 2022. [Accessed 12/07/2022]. Available from: https://www.simul8.com/support/help/doku.php?id=features:distributions:probability_profile

291. Szende, A., Janssen, M.F. and Cabases, J.M. *Self-Reported Population Health: An International Perspective based on EQ-5D.* 2014.
292. McDonald, L.C., Gerding, D.N., Johnson, S., Bakken, J.S., Carroll, K.C., Coffin, S.E., Dubberke, E.R., Garey, K.W., Gould, C.V., Kelly, C., Loo, V., Shaklee Sammons, J., Sandora, T.J. and Wilcox, M.H. Clinical Practice Guidelines for Clostridium difficile Infection in Adults and Children: 2017 Update by the Infectious Diseases Society of America (IDSA) and Society for Healthcare Epidemiology of America (SHEA). *Clin Infect Dis.* 2018, **66**(7), pp.e1-e48.
293. Lanzas, C., Dubberke, E.R., Lu, Z., Reske, K.A. and Gröhn, Y.T. Epidemiological Model for Clostridium difficile Transmission in Healthcare Settings. *Infect Control Hosp Epidemiol.* 2011, **32**(6), pp.553-561.
294. Planche, T. and Wilcox, M. Reference assays for *Clostridium difficile* infection: one or two gold standards? *J Clin Pathol.* 2011, **64**(1), p.1.
295. Techlab. C-DIFF-CHEK-60TM-60. [Online]. 2019. [Accessed 10/03/2021]. Available from: https://www.techlab.com/wp-content/uploads/2019/09/C-DIFF-CHEK-60_PI_91-392-03_read-version_08_2019.pdf
296. Cepheid. Xpert C.difficile ENGLISH Package Insert [Online]. 2019. [Accessed 10/03/2021]. Available from: https://p.widencdn.net/zmmjln/Cepheid-Xpert-C-difficile-BT-Datasheet-CE-IVD-3073-English
297. Cepheid. Xpert Cdifficile/Epi. [Online]. 2019. [Accessed 10/03/2021]. Available from: https://labymed.com.gt/wp-content/uploads/2020/06/XpertCdifficile-Epi.pdf
298. Department of Health. NHS reference costs 2014 to 2015. [Online]. 2012. [Accessed 18/11/22]. Available from: https://www.gov.uk/government/publications/nhs-reference-costs-2014-to-2015
299. National Services Scotland and health Protection Scotland. NHS Scotland MRSA Screening Pathfinder Progmame-Final Report Vlume 2: An Assessment of the Economics, Implementation and Modelling of Universital Screening. [Online]. 2011. [Accessed 18/11/22]. Available from: https://hps.scot.nhs.uk/web-resources-container/mrsa-screening-pathfinder-implementation-study-reports/
300. Sewell, B., Rees, E., Thomas, I., Ch'ng, C., Isaac, M. and Berry, N. Impact on patient length of stay and cost-effectiveness of rapid molecular testing for Clostridium difficile. *Clin Microbiol Infect.* 2012, **18**(SUPPL. 3), p.667.
301. British National Formulary. VANCOMYCIN Medicinal forms [Online]. [Accessed 20/05/2021]. Available from: https://bnf.nice.org.uk/medicinal-forms/vancomycin.html
302. Food Drug Administration Agency. Evaluation of Aumatic Class III Designation (De Novo) For xTAG® Gastrointestinal Pathogen Panel (GPP) Decision Summary. [Online]. 2012. [Accessed 09/02/2022]. Available from: www.accessdata.fda.gov/cdrh_docs/reviews/K121454.pdf
303. Buss, S., Leber, A., Chapin, K., Fey, P.D., Bankowski, M.J., Jones, M.K., Rogatcheva, M., Kanack, K.J., Bourzac, K.M. and Burnham, C.A.D. Multicenter Evaluation of the BioFire FilmArray Gastrointestinal Panel for

Etiologic Diagnosis of Infectious Gastroenteritis. *J Clin Microbiol.* 2015, **53**(3), pp.915-925.
304. Zheng, L., Keller, S.F., Lyerly, D.M., Carman, R.J., Genheimer, C.W., Gleaves, C.A., Kohlhepp, S.J., Young, S., Perez, S. and Ye, K. Multicenter Evaluation of a New Screening Test That Detects *Clostridium difficile* in Fecal Specimens. *J Clin Microbiol.* 2004, **42**(8), p.3837.
305. Bank of England Inflation Calculator. [Online]. [Accessed 05/03/2021]. Available from: https://www.bankofengland.co.uk/monetary-policy/inflation/inflation-calculator
306. Minor, T., Lasher, A., Klontz, K., Brown, B., Nardinelli, C. and Zorn, D. The Per Case and Total Annual Costs of Foodborne Illness in the United States. *Risk Anal.* 2015, **35**(6), pp.1125-1139.
307. Karnon, J., Stahl, J., Brennan, A., Caro, J.J., Mar, J. and Möller, J. Modeling using Discrete Event Simulation: A Report of the ISPOR-SMDM Modeling Good Research Practices Task Force-4. *Value in Health.* 2012, **15**(6), pp.821-827.
308. Sandlund, J., Davies, K. and Wilcox, M.H. Ultrasensitive *Clostridioides difficile* Toxin Testing for Higher Diagnostic Accuracy. *J Clin Microbiol.* 2020, **58**(6), pp.e01913-01919.
309. Hansen, G., Young, S., Wu Alan, H.B., Herding, E., Nordberg, V., Mills, R., Griego-Fullbright, C., Wagner, A., Ong Chui, M., Lewis, S., Yoon, J., Estis, J., Sandlund, J., Friedland, E., Carroll Karen, C. and Onderdonk Andrew, B. Ultrasensitive Detection of Clostridioides difficile Toxins in Stool by Use of Single-Molecule Counting Technology: Comparison with Detection of Free Toxin by Cell Culture Cytotoxicity Neutralization Assay. *J Clin Microbiol.* **57**(11), pp.e00719-00719.
310. Lee-Lewandrowski, E. and Lewandrowski, K. Perspectives on Cost and Outcomes for Point-of-Care Testing. *Clin Lab Med.* 2009, **29**(3), pp.479-489.
311. Barbut, F., Braun, M., Burghoffer, B., Lalande, V. and Eckert, C. Rapid detection of toxigenic strains of Clostridium difficile in diarrheal stools by real-time PCR. *J Clin Microbiol.* 2009, **47**(4), pp.1276-1277.
312. McCabe, C., Paulden, M., Awotwe, I., Sutton, A. and Hall, P. One-Way Sensitivity Analysis for Probabilistic Cost-Effectiveness Analysis: Conditional Expected Incremental Net Benefit. *PharmacoEconomics.* 2020, **38**(2), pp.135-141.
313. Korhonen, L., Cohen, J., Gregoricus, N., Farley, M.M., Perlmutter, R., Holzbauer, S.M., Dumyati, G., Beldavs, Z., Paulick, A., Vinjé, J., Limbago, B.M., Lessa, F.C. and Guh, A.Y. Evaluation of viral co-infections among patients with community-associated Clostridioides difficile infection. *PLOS ONE.* 2020, **15**(10), p.e0240549.
314. Shafiq, M., Alturkmani, H., Zafar, Y., Mittal, V., Lodhi, H., Ullah, W. and Brewer, J. Effects of co-infection on the clinical outcomes of Clostridium difficile infection. *Gut Pathog.* 2020, **12**(1), p.9.
315. Brown, K.A., Daneman, N., Arora, P., Moineddin, R. and Fisman, D.N. The Co-Seasonality of Pneumonia and Influenza With Clostridium difficile Infection in the United States, 1993–2008. *Am J Epidemiol.* 2013, **178**(1), pp.118-125.
316. Polgreen, P.M., Yang, M., Bohnett, L.C. and Cavanaugh, J.E. A Time-Series Analysis of Clostridium difficile and Its Seasonal Association with Influenza. *Infect Control Hosp Epidemiol.* 2010, **31**(4), pp.382-387.

317. Holmes, E.A.F. and Hughes, D.A. Challenges for Economic Evaluation of Health Care Strategies to Contain Antimicrobial Resistance. *Antibiotics.* 2019, **8**(4).
318. Rennert-May, E., Conly, J., Leal, J., Smith, S. and Manns, B. Economic evaluations and their use in infection prevention and control: a narrative review. *Antimicrob Resist Infect Control.* 2018, **7**(1), p.31.
319. Litvin, M., Reske, K.A., Mayfield, J., McMullen, K.M., Georgantopoulos, P., Copper, S., Hoppe-Bauer, J.E., Fraser, V.J., Warren, D.K. and Dubberke, E.R. Identification of a Pseudo-Outbreak of Clostridium difficile Infection (CDI) and the Effect of Repeated Testing, Sensitivity, and Specificity on Perceived Prevalence of CDI. *Infect Control Hosp Epidemiol.* 2009, **30**(12), pp.1166-1171.
320. Barker, A.K., Ngam, C., Musuuza, J.S., Vaughn, V.M. and Safdar, N. Reducing Clostridium difficile in the Inpatient Setting: A Systematic Review of the Adherence to and Effectiveness of C. difficile Prevention Bundles. *Infect Control Hosp Epidemiol.* 2017, **38**(6), pp.639-650.
321. Clinical & Laboratory Standards Institute. View Our Harmonized Terminology Database [Online]. 2021. [Accessed 23/05/2022]. Available from: https://clsi.org/standards-development/harmonized-terminology-database/#:~:text=The%20Harmonized%20Terminology%20Database%20is,and%20related%20health%20care%20industries.
322. Rahman, S., Majumder, M., Shaban, S.F., Rahman, N., Ahmed, M., Abdulrahman, K.B. and D'Souza, U.J. Physician participation in clinical research and trials: issues and approaches. *Adv Med Educ Pract.* 2011, (1179-7258 (Print)).
323. McRobert, C.J., Hill, J.C., Smale, T., Hay, E.M. and van der Windt, D.A. A multi-modal recruitment strategy using social media and internet-mediated methods to recruit a multidisciplinary, international sample of clinicians to an online research study. *PLOS ONE.* 2018, **13**(7), p.e0200184.
324. Noel, H.J. and Huang, A.R. The Effect of Varying Incentive Amounts on Physician Survey Response. *Eval Health Prof.* 2018, **42**(1), pp.71-81.
325. Cunningham, C.T., Quan, H., Hemmelgarn, B., Noseworthy, T., Beck, C.A., Dixon, E., Samuel, S., Ghali, W.A., Sykes, L.L. and Jetté, N. Exploring physician specialist response rates to web-based surveys. *BMC Med Res Methodol.* 2015, **15**(1), p.32.
326. Booker, Q.S., Austin, J.D. and Balasubramanian, B.A. Survey strategies to increase participant response rates in primary care research studies. *Fam Pract.* 2021, **38**(5), pp.699-702.
327. Creavin, S.T., Creavin, A.L. and Mallen, C.D. Do GPs respond to postal questionnaire surveys? A comprehensive review of primary care literature. *Fam Pract.* 2011, **28**(4), pp.461-467.
328. Ricci-Cabello, I., Avery, A.J., Reeves, D., Kadam, U.T. and Valderas, J.M. Measuring Patient Safety in Primary Care: The Development and Validation of the “Patient Reported Experiences and Outcomes of Safety in Primary Care” (PREOS-PC). *Ann Fam Med.* 2016, **14**(3), p.253.
329. Medina-Lara, A., Grigore, B., Lewis, R., Peters, J., Price, S., Landa, P., Robinson, S., Neal, R., Hamilton, W. and Spencer, A.E. Cancer diagnostic tools to aid decision-making in primary care: mixed-methods systematic reviews and cost-effectiveness analysis. *Health Technol Assess.* 2020, **24**, p.66.

330. Cho, Y.I., Johnson, T.P. and VanGeest, J.B. Enhancing Surveys of Health Care Professionals: A Meta-Analysis of Techniques to Improve Response. *Eval Health Prof.* 2013, **36**(3), pp.382-407.
331. VanGeest, J.B., Wynia, M.K., Cummins, D.S. and Wilson, I.B. Effects of Different Monetary Incentives on the Return Rate of a National Mail Survey of Physicians. *Med Care.* 2001, **39**(2).
332. Smith, A.F., Shinkins, B., Hall, P.S., Hulme, C.T. and Messenger, M.P. Toward a Framework for Outcome-Based Analytical Performance Specifications: A Methodology Review of Indirect Methods for Evaluating the Impact of Measurement Uncertainty on Clinical Outcomes. *Clin Chem.* 2019, **65**(11), pp.1363-1374.
333. Smith, A.F. Developing a framework for assessing the impact of test measurement uncertainty on clinical and health-economic outcomes. Doctoral thesis thesis, University of Leeds, 2020.
334. Borsci, S., Buckle, P., Uchegbu, I., Ni, M., Walne, S. and Hanna, G. Integrating human factors and health economics to inform the design of medical device: a conceptual framework. *IFMBE Proceedings.* 2018, pp.49-52.
335. Smith, R. and Schneider, P. *Data Science for Public Health Economic Evaluation.* [Online]. 2020. Available from: https://bookdown.org/rasmith3/r4phe2/
336. Mott, D.J., Hampson, G., Llewelyn, M.J., Mestre-Ferrandiz, J. and Hopkins, M.M. A Multinational European Study of Patient Preferences for Novel Diagnostics to Manage Antimicrobial Resistance. *Appl Health Econ Health Policy.* 2020, **18**(1), pp.69-79.
337. Graziadio, S., Urwin, S.G., Cocco, P., Micocci, M., Winter, A., Yang, Y., Price, D.A., Messenger, M., Allen, A.J., Shinkins, B. and on behalf of the, C.S.g. Unmet clinical needs for COVID-19 tests in UK health and social care settings. *PLOS ONE.* 2020, **15**(11), p.e0242125.
338. Sculpher, M., Fenwick, E. and Claxton, K. Assessing Quality in Decision Analytic Cost-Effectiveness Models. *PharmacoEconomics.* 2000, **17**(5), pp.461-477.
339. Siebert, U., Alagoz, O., Bayoumi, A.M., Jahn, B., Owens, D.K., Cohen, D.J. and Kuntz, K.M. State-Transition Modeling: A Report of the ISPOR-SMDM Modeling Good Research Practices Task Force-3. *Value in Health.* 2012, **15**(6), pp.812-820.
340. Kim, S.-Y. and Goldie, S.J. Cost-Effectiveness Analyses of Vaccination Programmes. *PharmacoEconomics.* 2008, **26**(3), pp.191-215.
341. Viana, J., Brailsford, S.C., Harindra, V. and Harper, P.R. Combining discrete-event simulation and system dynamics in a healthcare setting: A composite model for Chlamydia infection. *Eur J Oper Res.* 2014, **237**(1), pp.196-206.
342. Chahal, K., Eldabi, T. and Young, T. A conceptual framework for hybrid system dynamics and discrete event simulation for healthcare. *J Enterp Inf Manag.* 2013, **26**(1/2), pp.50-74.
343. Caro, J.J., Möller, J., Karnon, J., Stahl, J. and Ishak, J. *Discrete Event Simulation for Health Technology Assessment.* Chapman and Hall/CRC, 2016.
344. MedicinePlus Medical Encyclopedia and National Library of Medicine. Anaerobic. [Online]. 2020. [Accessed 01/06/22]. Available from: https://medlineplus.gov/ency/article/002230.htm

345. Medicine Plus Medical Encyclopedia and National Library of Medicine. Antigen. [Online]. 2020. [Accessed 01/06/22]. Available from: https://medlineplus.gov/ency/article/002224.htm
346. Medicine Plus Medical Encyclopedia and National Library of Medicine. Antibody. [Online]. 2018. [Accessed 01/06/22]. Available from: https://medlineplus.gov/ency/article/002223.htm#:~:text=An%20antibody%20is%20a%20protein,%2C%20and%20viruses)%20and%20chemicals.
347. National Cancer Institute Dictionary of Cancer Terms Definition of intestinal flora. [Online]. 2022. [Accessed 01/06/22]. Available from: https://www.cancer.gov/publications/dictionaries/cancer-terms/def/intestinal-flora
348. Medicine Plus Medical Encyclopedia and National Library of Medicine. Campylobacter infection. [Online]. 2020. [Accessed 01/06/22]. Available from: https://medlineplus.gov/ency/article/000224.htm
349. Medicine Plus Medical Encyclopedia and National Library of Medicine. Celiac Disease  Gluten Intolerance [Online]. 2022. [Accessed 01/06/22]. Available from: https://medlineplus.gov/celiacdisease.html
350. Medicine Plus Medical Encyclopedia and National Library of Medicine. Colitis. [Online]. 2016. [Accessed 01/06/22]. Available from: https://medlineplus.gov/ency/article/001125.htm
351. Dani, A. Colonization and infection. *Central European journal of urology.* 2014, **67**(1), pp.86-87.
352. Medicine Plus Medical Encyclopedia and National Library of Medicine. Creatinine Test. [Online]. 2020. [Accessed 01/06/22]. Available from: https://medlineplus.gov/lab-tests/creatinine-test/
353. National Cancer Institute Dictionary of Cancer Terms Definition of enteral nutrition [Online]. 2022. [Accessed]. Available from: https://www.cancer.gov/publications/dictionaries/cancer-terms/def/enteral-nutrition
354. Medicine Plus Medical Encyclopedia and National Library of Medicine. E. Coli Infections. [Online]. 2016. [Accessed 01/06/22]. Available from: https://medlineplus.gov/ecoliinfections.html
355. Bustillos, H., Leer, K., Kitten, A. and Reveles, K.R. A cross-sectional study of national outpatient gastric acid suppressant prescribing in the United States between 2009 and 2015. *PLOS ONE.* 2018, **13**(11), pp.e0208461-e0208461.
356. National Cancer Institute Dictionary of Cancer Terms Definition of stomach cancer. [Online]. 2022. [Accessed 01/06/22]. Available from: https://www.cancer.gov/publications/dictionaries/cancer-terms/def/stomach-cancer
357. National Cancer Institute Dictionary of Cancer Terms Definition of neoplasm [Online]. 2022. [Accessed 01/06/22]. Available from: https://www.cancer.gov/publications/dictionaries/cancer-terms/def/neoplasm
358. Medicine Plus Medical Encyclopedia and National Library of Medicine. Gram Stain. [Online]. 2020. [Accessed 01/06/22]. Available from: https://medlineplus.gov/lab-tests/gram-stain/
359. Medicine Plus Medical Encyclopedia and National Library of Medicine. Low blood pressure. [Online]. 2019. [Accessed 01/06/22]. Available from: https://medlineplus.gov/ency/article/007278.htm
360. Beach, E.C. and De Jesus, O. Ileus. [Online]. 2021. [Accessed 01/06/22]. Available from: https://www.ncbi.nlm.nih.gov/books/NBK558937/

361. National Cancer Institute Dictionary of Cancer Terms Definition of inflammatory bowel disease. [Online]. 2022. [Accessed 01/06/22]. Available from: https://www.cancer.gov/publications/dictionaries/cancer-terms/def/inflammatory-bowel-disease
362. Medicine Plus Medical Encyclopedia and National Library of Medicine. Irritable Bowel Syndrome. [Online]. 2018. [Accessed 01/06/22]. Available from: https://medlineplus.gov/irritablebowelsyndrome.html
363. National Cancer Institute Dictionary of Cancer Terms Definition of in vitro. [Online]. 2022. [Accessed 01/06/22]. Available from: https://www.cancer.gov/publications/dictionaries/cancer-terms/def/in-vitro
364. National Cancer Institute Dictionary of Cancer Terms Definition of multiplex genomic test. [Online]. 2022. [Accessed 01/06/2022]. Available from: https://www.cancer.gov/publications/dictionaries/genetics-dictionary/def/multiplex-genomic-test
365. NHS. Urinary incontinence [Online]. 2019. [Accessed 01/06/22]. Available from: https://www.nhs.uk/conditions/urinary-incontinence/
366. Merriam-Webster.com Medical Dictionary. Oxidative stress Definition & Meaning [Online]. 2022. [Accessed 01/06/22]. Available from: https://www.merriam-webster.com/medical/oxidative%20stress
367. Medicine Plus Medical Encyclopedia and National Library of Medicine. Salmonella Infections. [Online]. 2016. [Accessed 01/06/22]. Available from: https://medlineplus.gov/salmonellainfections.html
368. Medline Plus Medical Encyclopedia and National Library of Medicine. Shigellosis. [Online]. 2020. [Accessed 01/06/22]. Available from: https://medlineplus.gov/ency/article/000295.htm
369. Medicine Plus Medical Encyclopedia and National Library of Medicine. Spores. [Online]. 2019. [Accessed 01/06/22]. Available from: https://medlineplus.gov/ency/article/002307.htm#:~:text=A%20spore%20is%20a%20cell,Spores%20have%20thick%20walls.
370. Biology Online Dictionary. Strain Definition and Examples. [Online]. 2021. [Accessed 01/06/22]. Available from: https://www.biologyonline.com/dictionary/strain
371. Skomorochow, E. and Pico, J. Toxic Megacolon. BTI - StatPearls. [Online]. 2022. [Accessed 01/06/22]. Available from: https://www.ncbi.nlm.nih.gov/books/NBK547679/
372. Medicine Plus Medical Encyclopedia and National Library of Medicine. White Blood Count (WBC). [Online]. 2020. [Accessed 01/06/22]. Available from: https://medlineplus.gov/lab-tests/white-blood-count-wbc/#:~:text=A%20white%20blood%20count%20measures,foreign%20substances%20causing%20your%20illness.
373. DiPersio, J.R., Varga, F.J., Conwell, D.L., Kraft, J.A., Kozak, K.J. and Willis, D.H. Development of a rapid enzyme immunoassay for Clostridium difficile toxin A and its use in the diagnosis of C. difficile-associated disease. *J Clin Microbiol.* 1991, (0095-1137 (Print)).

# Appendix A
# Overview of modelling techniques

This appendix describes the main features, advantages and limitations of five modelling techniques that are relevant for the purposes of this thesis (e.g. decision tree models, Markov models, system dynamics models, and discrete event simulation models, agent-based simulation models).

## A.1 Decision trees models

A decision tree illustrates the possible prognoses and/or events a cohort of individuals can experience due to a certain healthcare intervention, as depicted in a sequence of pathways (61). Decision trees typically feature key elements, including: (i) *decision nodes* – which represent the decisions being assessed; (ii) *chance nodes* – which illustrate a range of possible subsequent events conditional on a certain decision; (iii) *branches* – which represent the link between the decision nodes and the subsequent chance nodes conditional of taking a certain decision; (iv) *pathways* – which encompasses the combination of different branches following a certain decision. The branches represent mutually exclusive pathways individuals can experience following a certain decision; and (v) *terminal nodes* – which are positioned at the end of each pathway and represent the outcome being evaluated (e.g. costs, QALYs).

Each branch is associated with certain costs and probabilities. Total pathway costs include the costs associated with each of the events the cohort experiences, whilst pathway probabilities represent the product of the probabilities in that given pathway (61).

Decision trees are characterised by a simple structure, which is usually recommended while developing any decision model (273, 338), alongside a short time-horizon which is suitable for mapping diagnostic pathways (274). Depending on the decision being evaluated and the mechanisms a modeller wish to include, there are disadvantages in relation to decision trees. This modelling technique simulates individuals at the aggregate level, therefore individual characteristics (e.g. age, risk factors) are not captured (273). Decision trees also assume no interactions among individuals (273), therefore these models are not recommended to simulate interactions between individuals, key resources and processes within a clinical pathway, as well as spread of infection. Decision trees

also do not feature looping events, so individuals cannot re-enter the model once they have reached a given terminal point (273). In addition to the above, it is not possible to explicitly simulate the passage of time within decision trees as time is assumed to be fixed (273, 274).

## A.2 Markov models

A Markov model is represented by a sequence of mutually exclusive health states in which a patient (or group of patients) can remain or move in between at a certain point in time (61). The probability of moving between health states is represented by transition probabilities. At each discrete unit of time (i.e. *cycle*), patients can either stay in the same state or move to progressive disease state or die, depending on the assigned transition probabilities. A key assumption of Markov models is the *memoryless assumption* which means that the transition probabilities in a certain time point are independent from the earlier transitions between the different health states (61). Within Markov models, the passage of time is measured deterministically[29] based on an arbitrary cycle length (e.g. 1 month) during which patients are allowed to move across the health states (273, 274). Costs and health-related utilities are associated to each health state. To compute the total expected costs and effects, the number of patients in a certain health state at a given time point is weighted to the costs and utilities for that health state.

There are two types of Markov models: cohort-level and individual-level Markov models. Whilst cohort-level Markov models can capture recurring events, individual-level Markov models are better suited to this task as they are able to account for the impact of individual patient histories on future transition probabilities (339). In the context of cohort-level Markov models, a key assumption is *homogeneity* and *independency* between individuals (273) – therefore individual characteristics are not captured, and it is not possible to simulate interactions between individuals, infection spread or capacity constraints (271-273). As transition probabilities are independent on the patient's history and patients can only remain in one health state each cycle, many health states may

[29] It is possible to build Markov models where the passage of time is measured continuously, rather than at discrete time points (275).

be required to fully represent the disease progression thereby leading to an excessive number of health states.

## A.3 System dynamics models

A SD model is a cohort-level modelling technique capturing interactions between individuals within a certain system (273). This modelling technique stratifies the population into subgroups of compartments, and groups of individuals move between compartments based on key parameters (340). In the context of infectious diseases, common compartments usually entail individuals who are susceptible, colonised, infectious and recovered. An additional feature of SD models is that they are considered memoryless – they do not account for individuals' past history or movements.

Within SD models, the passage of time is continuous as the simulation clock does not stop to evaluate mid-states outcomes and events – rather this modelling technique uses rates and/or equations to evaluate changes over time. This means that SD models allow for constant updating of model parameters based on changes of the system itself (e.g. number of individuals within each compartments). For example, the infection rate iteratively changes based on the number of susceptible and infected patients (279).

While SD models take a holistic perspective exploring the aggregate consequences of changes in the clinical pathway on the wider hospital system (279, 341, 342), they are not able to capture capacity constraints or key processes in place in a given health care setting (279) – as opposed to DES and ABS models.

## A.4 Discrete event simulation models

DES is an individual-based modelling technique which captures the progression of single passive individuals (i.e. *entities*) through different processes and activities (i.e. *events*) (273). Different unique characteristics (i.e. *attributes*) – such as age, duration of symptoms and risk of experiencing a certain event – can be assigned to each patient which, in turn, will affect the natural disease progression, the future events and outcomes a given individual might experience (271, 273, 274). As a result, patient history or memory can be captured within DES models (275). DES models can also simulate interactions between individuals within a given system (273, 275). In addition to this, DES models are usually stochastic

as they are subject to *first-order* uncertainty which reflects "the fact that individuals facing the same probabilities and outcomes will experience the effects of a disease or intervention differently [..] (e.g. the first patient in a sample might respond to a treatment but the next one may not)" (280).

Within DES models, transitions between events are based on probabilities and they are evaluated as they occur rather than at fixed time points (as per Markov models). Time advances at discrete times depending on when the next event is scheduled to occur (i.e. *next time-to-event progression*) (273).

Selecting a model which accounts for patient history and/or past patient transitions is helpful to determine the future patient trajectory within the clinical pathway, disease progression and associated resource requirements based on patient's current status (e.g. test result, infectious status). This enables evaluation of capacity implications as it is possible to account for what resources each individual patient needs (e.g. single rooms).

DES models can track the individual journey patients experience within a certain healthcare system, so they can capture resource constraints[30] via implementing logic of queues (343). In case where a given resource (e.g. isolation rooms) is occupied, individuals will be required to wait until the resource is available according to a pre-specified logic (e.g. *first-in-first-out*, *last in-first-out*) (279).

As per every modelling technique, DES models presents some limitations. They require extensive data to inform patient attributes, are more computationally intensive and stochastic as individual patients are tracked during the simulation using a random process (273, 274).

In relation to modelling infectious diseases, despite DES models allowing for interactions between entities and patients, the time-to-next event progression and individual features of this technique make the task of updating the number of infectious and susceptible individuals (in order to estimate the current infectious rate) computationally intensive (279). DES models cannot efficiently update key model parameters (e.g. rate of infection) as a result of changes in model outputs (e.g. number of colonised and infected patients) – as opposed to SD models (279). In addition to this, the stochastic element of DES models reduces the

[30] It is possible, however, to develop non constrained-resource DES models where the impact of queues and limited resources is not captured, and the availability of resources is assumed to be infinite.

applicability of this modelling technique in the context of complex infectious disease modelling.

## A.5 Agent-based simulation models

ABS models are individual-based simulations capturing the interaction between autonomous individual entities (i.e. *agents*) and the simulated system. The focus of this modelling technique lies in capturing the agents' predicted behaviours, goals and decisions rules and how they interact and sense the system (279).

Within ABM models passage of time can be either discrete or continuous (343). Compared to DES models, there are many similarities, including: (i) different attributes can be assigned to each entity which, in turn, will affect the future progression within the system and future events; (ii) the ability to capture patient history or memory; (iii) the ability to capture resource constraints via implementing logic of queues'; and (iv) the ability to simulate interactions between entities.

Similar to DES models, ABS models require extensive and detailed data on the agents' attributes and predicted behaviour for parameterisation and validation purposes (279). In addition, this modelling technique is computationally intensive and conducting extensive sensitivity analysis appears to be challenging (279).

# Appendix B

# Search strategies – systematic review of TPPs for medical tests

This appendix provides details on the database (section B.1) and website search strategies (section B.2) for the systematic review of TPPs for medical tests presented in Chapter 2.

## B.1. Database search strategy

**Table B-1 Search strategy for database search**

| # | Keywords |
|---|---|
| 10 | #8 OR #9 |
| 9 | (TPP adj5 (test* OR assay OR screen* OR exam* OR diagnos* OR analys#s)).ti,ab,kw |
| 8 | #4 AND #7 |
| 7 | #5 OR #6 |
| 6 | (test* OR assay OR screen* OR exam* OR diagnos* OR analys#s).ti,ab,kw |
| 5 | exp "Diagnostic Techniques and Procedures"/ |
| 4 | #1 OR #2 OR #3 |
| 3 | “quality by design”.ti,ab,kw |
| 2 | QTPP.ti,ab,kw |
| 1 | “target product profil*”.ti,ab,kw |

**Table B-2 List of databases searched and number of references retrieved**

| Database searched | Date of search | Total number of results found (n) |
|---|---|---|
| Ovid MEDLINE ® (1946 to October Week 5 2018) | 12/11/18 | 784 |
| Ovid MEDLINE(R) In-Process & Other Non-Indexed Citations November 09, 2018 | 12/11/18 | 108 |
| Ovid MEDLINE(R) and Epub Ahead of Print (November 09 2018) | 12/11/18 | 24 |
| Embase Classic + Embase 1947 | 12/11/18 | 1896 |
| CAB Abstract online 1910 to 2018 Week 44 | 12/11/18 | 125 |
| Global Health (1910 to 2018 Week 44) | 12/11/18 | 102 |
| CINAHL | 12/11/18 | 237 |
| Scopus | 12/11/18 | 1396 |
| Web of Science Core Collection | 12/11/18 | 1454 |

## B.2. Website search

The structured methods proposed by Godin, K. et al. (92) were adopted to conduct the website search.

### B.2.1. Search for relevant websites

**Table B-3 Search for relevant website results**

| Number of research | Date | Search Engine | Keywords | # new potentially relevant records | Total # records | Notes |
|---|---|---|---|---|---|---|
| 1 | 18/10/2018 | Google | "target product profile" AND diagnostic test | 16 | 16 | Only 350 results found |
| 2 | 18/10/2018 | Google | "target product profile" AND screening | 3 | 19 | Only 308 results found |
| 3 | 18/10/2018 | Google | "target product profile" AND exam OR test OR assay | 4 | 24 | Only 304 results found |
| 4 | 18/10/2018 | Google | TPP AND diagnostic test | 0 | 24 | Only 245 results found |
| 5 | 18/10/2018 | Google | TPP AND screening | 0 | 24 | Only 332 results found |
| 6 | 18/10/2018 | Google | TPP AND exam OR test OR assay | 0 | 24 | Only 364 results found |
| 7 | 18/10/2018 | Google | QTPP AND diagnostic test | 2 | 26 | Only 119 results found |
| 8 | 18/10/2018 | Google | QTPP AND screening | 0 | 26 | Only 307 results found |
| 9 | 18/10/2018 | Google | QTPP AND exam OR test OR assay | 0 | 26 | Only 290 results found |

### B.2.2. List relevant websites

**Table B-4 Detailed results of search for relevant organisations and related websites**

| # search | Name organization | URL |
|---|---|---|
| 1 | FIND | https://www.finddx.org |
| | WHO | http://www.who.int |
| | The International Diagnostics centre London School of Hygiene & Tropical Medicine | http://www.idc-dx.org |
| | UNICEF | https://www.unicef.org |
| | PATH | https://path.org |
| | DNDi | https://www.dndi.org/ |
| | FDA | https://www.fda.gov |
| | Pink Ribbon | http://pinkribbonredribbon.org/ |
| | Unitaid | https://unitaid.org |
| | Malaria consortium | https://www.malariaconsortium.org/ |
| | ReAct group | https://www.reactgroup.org/ |
| | DiaDev (Investigating diagnostics in global health) | http://www.diadev.eu/ |
| | Public Health Emergency | https://www.phe.gov/ |
| | Galvmed | https://www.galvmed.org |
| | STOP TB Partnership | http://www.stoptb.org |
| 2 | EMA (European Medicines Agency) | https://www.ema.europa.eu/ |
| | NICE | https://www.nice.org.uk/ |
| | IVCC | http://www.ivcc.com/ |
| 3 | Critical Path Institute | https://c-path.org/ |
| 4 | European Directorate for the Quality of Medicines & HealthCare | https://www.edqm.eu/ |
| 5 | AIGHD | https://www.aighd.org/ |
| 6 | IMPT for Reproductive Health | https://www.theimpt.org/ |
| 7 | ICH harmonisation for better health | https://www.ich.org |
| | Product Quality Research Institute | http://pqri.org/ |
| 8 | N/A | N/A |
| 9 | N/A | N/A |

### B.2.3. Results website search

**Keywords:** “target product profile” in Title

**Included**: documents and publications relevant for the research question based on the title

**Excluded**: PowerPoint presentation, newsletters and results not relevant for the research question based on the title

**Date search**: 18-22/10/18

## Table B-5 Details on website search and number of relevant and included references

| Date | Website name | URL | Specific section searched | Search terms used | Total number of hits | Approach to screening for relevancy | Number of relevant references | Notes related to screening | Number of included references | Notes related to inclusion |
|---|---|---|---|---|---|---|---|---|---|---|
| 18/10/2018 | FIND | https://www.finddx.org | Internal search engine | "target product profile" | 76 | Title | 19 | | **22** | The number of included references entails relevant hits that were retrieved from both searches within the same website. Duplicates across both searches were deleted. |
| 18/10/2018 | | https://www.finddx.org/target-product-profiles/ | Hand-searched in the section of the FIND website dedicated to target product profiles | "target product profile" | | Title | 14 | | | |
| 18/10/2018 | WHO | http://www.who.int/ | Internal search engine | "target product profile" | 19 | Title | 11 | Search for only for publications | **28** | The number of included references entails relevant hits that were retrieved from both searches within the same website. Duplicates across both searches were deleted. |
| 18/10/2018 | | http://apps.who.int/iris | Internal database within the WHO website | "target product profile" | 29 | Title | 17 | Search for articles "meeting Abstracts" | | |
| 18/10/2018 | PATH | https://www.path.org | Internal search engine | "target product profile" | 228 | Title | 24 | | **24** | |
| 18/10/2018 | The International Diagnostics centre London School of Hygiene & Tropical Medicine | http://www.idc-dx.org/ | Internal search engine | "target product profile" | 8 | Title | 8 | Search for - publication - guidelines - report -white paper | **8** | |
| 18/10/2018 | | http://www.idc-dx.org/themes/development-and-evaluation/target-product-profiles | Hand-searched in the section of the website dedicated to target product profiles | "target product profile" | | Title | 6 | | | |

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| 18/10/2018 | UNICEF | https://www.unicef.org | Internal search engine | "target product profile" | 67 | Title | 7 | Search for document | **7** | The number of included references entails relevant hits that were retrieved from both searches within the same website. Duplicates across both searches were deleted. |
| 18/10/2018 | | https://www.unicef.org/publications/index_search.php | Search into publication section | "target product profile" | 0 | | 7 | | | |
| 18/10/2018 | DNDi | https://www.dndi.org | Internal search engine | "target product profile" | 27 | Title | 6 | | **6** | The number of included references entails relevant hits that were retrieved from both searches within the same website. Duplicates across both searches were deleted. |
| 18/10/2018 | | https://www.dndi.org/diseases-projects/target-product-profiles/ | Hand-searched in the section of the website dedicated to target product profiles | "target product profile" | | Title | 4 | | | |
| 18/10/2018 | FDA | https://www.fda.gov | Internal search engine | "target product profile" | 69 | Title | | | **0** | |
| 18/10/2018 | Pink ribbon | http://pinkribbonredribbon.org | Internal search engine | "target product profile" | 0 | Title | 0 | | **0** | |
| 22/10/2018 | Unitaid | https://unitaid.org | Internal search engine | "target product profile" | 2 | Title | 1 | I only found newsletters and a potentially relevant article | **1** | |
| 22/10/2018 | Malaria Consortium | https://www.malariaconsortium.org/ | Internal search engine | "target product profile" | 15 | Title | 4 | Powerpoint presentations and newsletters were excluded | **4** | |
| 22/10/2018 | ReAct Group | https://www.reactgroup.org | Internal search engine | "target product profile" | 2 | Title | 0 | Records in the "News & Views" section were excluded | **0** | |
| 22/10/2018 | DiaDev | http://www.diadev.eu | Internal search engine: "on the web" section | "target product profile" | 0 | Title | 0 | There is no internal search engine | | |
| 22/10/2018 | | http://www.diadev.eu | Internal search engine: "presentation" section | "target product profile" | 0 | Title | 0 | | | |

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | | | **0** | |
| 22/10/2018 | | http://www.diadev.eu | Internal search engine: "academic publications" section | "target product profile" | 0 | Title | 0 | | | |
| 22/10/2018 | | http://www.diadev.eu | Internal search engine "working papers" section | "target product profile" | N/A | Title | N/A | I could not access this section. Blank page | | |
| 22/10/2018 | | http://www.diadev.eu | Hand-searched in the "connections" section | "target product profile" | 0 | Title | 0 | | | |
| 22/10/2018 | | http://www.diadev.eu | Hand-searched "methods" section | "target product profile" | 2 | Title | 0 | | | |
| 22/10/2018 | | http://www.diadev.eu | Hand-searched "device case studies" section | "target product profile" | 0 | Title | 0 | | **0** | |
| 22/10/2018 | Public Health Emergency | https://www.phe.gov/ | Internal search engine | "target product profile" | N/A | Title | N/A | The internal search engine is not functioning | **N/A** | |
| 22/10/2018 | Galvmed | https://www.galvmed.org | Internal search engine | "target product profile" | 3 | Title | 0 | Reason for not including the records: TPP for animals | **0** | |
| 22/10/2018 | | https://www.galvmed.org/resources/databases/document-repository/ | Internal search into Document Repository database | "target product profile" | 0 | Title | 0 | | | |
| 22/10/2018 | | https://www.galvmed.org/resources/databases/livestock-laws-and-policy-database/ | Internal search into Livestock Livestock and Policy Database | "target product profile" | 119 | Title | 0 | | **0** | |
| 22/10/2018 | STP TB Partnership | http://www.stoptb.org | Internal search engine | "target product profile" | 48 | Title | 0 | Powerpoint presentations and newsletters were excluded *Google Custom Search | | The number of included references entails relevant hits that were retrieved from both |

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| 22/10/2018 | | http://www.stoptb.org | Hand-searched in "Technical Publications" section | "target product profile" | | Title | 1 | | **1** | searches within the same website. Duplicates across both searches were deleted. |
| 22/10/2018 | EMA | https://www.ema.europa.eu | Internal search engine | "target product profile" | 210 | Title | 0 | Search only for documents | **0** | |
| 22/10/2018 | NICE | https://www.nice.org.uk | Internal search engine | "target product profile" | 0 | Title | 0 | | **0** | |
| 22/10/2018 | | https://www.nice.org.uk | Internal search into the "Evidence search" engine | "target product profile" | 9 | Title | 0 | | **0** | |
| 22/10/2018 | IVCC | http://www.ivcc.com | Internal search engine | "target product profile" | 4 | Title | 0 | Only newsletters found | **0** | |
| 22/10/2018 | Critical Path Institute | https://c-path.org | Internal search engine | "target product profile" | 3 | Title | 1 | | **1** | |
| 22/10/2018 | European Directorate for the Quality of Medicines & HealthCare | https://www.edqm.eu/ | Internal search engine | "target product profile" | 5 | Title | 0 | | **0** | |
| 22/10/2018 | AIGHD | https://www.aighd.org/ | Search into the "Scientific Publications & PhD theses" | "target product profile" | N/A | Title | N/A | Blank page as a result of the search | | |
| 22/10/2018 | | https://www.aighd.org/ | Internal search engine | "target product profile" | 0 | Title | N/A | | **0** | |
| 22/10/2018 | IMPT for Reproductive Health | https://www.theimpt.org/ | Search into "Resource Database" | "target product profile" | 3 | Title | 3 | | | |
| 22/10/2018 | | https://www.theimpt.org/ | Internal search engine | "target product profile" | 11 | Title | 4 | | **4** | The number of included references entails relevant hits that were retrieved from both searches within the same website. Duplicates across both searches were deleted. |
| 22/10/2018 | ICH harmonisa | https://www.ich.org | Internal search engine | "target product profile" | 0 | Title | 0 | | **0** | |

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | tion for better health | | | | | | | | | |
| 22/10/2018 | Product Quality Research Institute | http://pqri.org/ | Internal search engine | "target product profile" | 0 | Title | 0 | | | |
| 22/10/2018 | | http://pqri.org/ | Search into "White papers" section | "target product profile" | 0 | Title | 0 | | **0** | |

# Appendix C

# Additional results – systematic review of TPPs

This appendix provides additional results from the systematic review of TPPs for medical tests presented in Chapter 2, including (i) calculation of inter-reviewer agreement rate (see section C.1); (ii) full data extraction table (section C.2); and (iii) results from transparency assessment of the included TPPs (section C.3).

## C.1. Calculation inter-reviewer agreement rate (κ statistic)

### Title/Abstract screening

**Table C-1 Title/Abstract screening inter-reviewer agreement rate calculation**

| | | BS | | |
|---|---|---|---|---|
| | | Included | Excluded | Tot |
| PC | Included | 13 | 1 | 14 |
| | Excluded | 0 | 275 | 275 |
| | Tot | 13 | 276 | 291 |

PC – Paola Cocco; BS – Bethany Shinkins

$$p_e = [({}^{n_1}/_{n}) * ({}^{m_1}/_{n})] + [({}^{n_0}/_{n}) * ({}^{m_0}/_{n})]$$
$$= [({}^{13}/_{289}) * ({}^{14}/_{289})] + \left[({}^{276}/_{289}) * \left({}^{275}/_{289}\right)\right] = 91\%$$

$$p_0 = \frac{(a+d)}{n} = \frac{(13+275)}{289} = 100\%$$

$$\kappa = \frac{(p_0 - p_e)}{(1-p_e)} = \frac{(100\% - 91\%)}{(1-91\%)} = 96\%$$

### Full text screening

**Table C-2 Full text screening inter-reviewer agreement rate calculation**

| | | BS | | |
|---|---|---|---|---|
| | | Included | Excluded | Tot |
| PC | Included | 44 | 0 | 44 |
| | Excluded | 1 | 63 | 64 |
| | Tot | 45 | 63 | 108 |

$$p_e = [({}^{n_1}/_{n}) * ({}^{m_1}/_{n})] + [({}^{n_0}/_{n}) * ({}^{m_0}/_{n})]$$
$$= \left[\left({}^{45}/_{108}\right) * ({}^{44}/_{108})\right] + [({}^{63}/_{108}) * ({}^{64}/_{108})] = 91\%$$

$$p_0 = \frac{(a+d)}{n} = \frac{(44+63)}{108} = 99\%$$

$$\kappa = \frac{(p_0 - p_e)}{(1-p_e)} = \frac{(99\% - 91\%)}{(1-91\%)} = 90\%$$

## C.2. Data extraction

**Table C-3 Extended summary table data extraction: publication format, disease area, funding body, test, aim TPP, decision-making phases described, description of decision-making process, input sources) to inform TPP and stakeholder involved for each included TPP (n=44)**

| Authors | Publication format | Disease area | Funding body | Test(s) | Aim | Decision-making phases described | Decision-making process | Input source(s) to inform TPP[31] | Stakeholder(s) involved |
|---|---|---|---|---|---|---|---|---|---|
| **Chua, A. et al. (98)** | Journal article | Zika virus | /[32] | Diagnostic test for Zika infection and blood bank testing | / | -Drafting,<br>-Consensus-building | -Meeting with experts,<br>-Draft TPP,<br>-Consensus meeting,<br>-Revision following feedback | Meeting input (D),<br>Literature (D) | Researchers (D),<br>International public organisation(D)<br>Scientific associations (D) |
| **Denkinger, C. et al. (99)** | Journal article | Tubercolosis | Bill and Melinda Gates Foundation; American Society of Tropical Medicine and Hygiene; National Institute of Allergy and Infectious Diseases, National Institutes of Health, Department of Health and Human Services; | A molecular drug-susceptibility test | -To define minimal test characteristic as lowest acceptable specification<br>- To define optimal test characteristic as ideal value for that characteristic | -Scoping,<br>-Drafting<br>-Consensus-building | -Mapping,<br>-Survey to measure stakeholders' preferences,<br>-Landscaping exercise,<br>-Draft TPP,<br>-Round of revisions,<br>-Shortened TPP draft,<br>-Presentation of shortened TPP to stakeholders,<br>-Consensus meeting | Literature (S,D, diagnostic accuracy[33])<br>Expert opinion (S,D),<br>Models (D, diagnostic accuracy),<br>Data available (S,D),<br>Market analyses (S) | Researchers (S),<br>Clinicians (S,D,C),<br>Policy makers (C),<br>Industry representatives (S),<br>Laboratory experts (S,D),<br>Implementers (C),<br>Representatives of national disease programs (C) |

[31] S= scoping phase; D= Drafting phase ; C= Consensus-building phase
[32] / = No information
[33] Input source used to inform diagnostic accuracy

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| **Denkinger, C. et al. (100)** | Journal article | Tubercolosis | Bill and Melinda Gates Foundation | 1) A diagnostic test<br>2) A diagnostic biomarker test<br>3) A referral screening test | / | -Scoping<br>-Drafting<br>-Consensus-building | -Definition problem statement,<br>-Priority-setting,<br>- Draft TPP,<br>-Round of revisions,<br>-Shortened draft TPP,<br>-Presentation of shortened TPP to stakeholders,<br>-Delphi-like survey to gauge stakeholders' agreement with the TPP,<br>-Consensus meeting | Literature (S,D, diagnostic accuracy),<br>Expert opinion (S,D)<br>Reports (S)<br>Models (D, diagnostic accuracy) | Representatives of national disease programs (S,C),<br>Clinicians (S,D,C),<br>Researchers (S,D),<br>Laboratory experts (S),<br>Patient advocates (S),<br>Modelers (S),<br>Market experts (S),<br>Policy makers (D),<br>Industry representatives (D,C),<br>Technical/funding agencies (D) |
| **DIAMETER Project and PATH. (101)** | Report | Malaria | Bill and Melinda Gates Foundation | A POC infection detection test | / | / | / | Literature (D, analytical accuracy[34]),<br>Available data (D),<br>Field observation (D),<br>Models (D),<br>Expert opinion (analytical accuracy) | / |
| **Ding, X. et al. (102)** | Journal article | Malaria | Department of Foreign Affairs and Trade, Australia | 1) A diagnostic test<br>2) A POC diagnostic test<br>3) A screening test | -To define minimal test characteristic as a value set to provide a distinguishing advantage<br>- To define minimal test characteristic as a value that provides optimal diagnostic effectiveness | -Scoping,<br>-Drafting<br>-Consensus-building | -Definition problem statement,<br>-Draft TPP,<br>-Round of revisions TPP,<br>-Survey to measure stakeholders' preferences | Literature (D, analytical accuracy),<br>Meeting input (S),<br>Expert opinion (S,D)<br>Models (D,analytical accuracy) | Researchers (S,C),<br>Representatives of national disease programs (S,C),<br>International public organisation (S,C) |

[34] Input source used to inform analytical accuracy

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| **Dittrich, S. et al. (103)** | Journal article | Malaria | The Dutch Government, with aid from the UK and the Australian Government | A test to distinguish bacterial from non-bacterial infections | -To define minimal test characteristic as lowest acceptable specification<br>- To define optimal test characteristic as ideal value for that characteristic | -Scoping,<br>-Drafting,<br>-Consensus-building | -Landscaping exercise,<br>-Draft TPP,<br>-Delphi-like approach to gauge stakeholders' agreement with the TPP,<br>-Round of revisions,<br>-Revision following feedback ,<br>-Priority-setting,<br>-Consensus meeting | Literature (S,D),<br>Expert opinion (D, analytical accuracy),<br>Industry standard (D) | Researchers (D,C),<br>Laboratory experts (D,C) ,<br>Microbiologists (D,C),<br>Health economists (D,C),<br>Industry representatives (D,C),<br>International public organisation (D,C) |
| **Donadeu, M. et al. (104)** | Journal article | Taenia solium taeniasis, neurocysticercosis and porcine cysticercosis | / | 1) A POC test that could be used for surveillance<br>2) A specific test<br>3)A POC test<br>4) A monitoring test | / | -Scoping,<br>-Drafting,<br>-Consensus-building | -Introductory meeting<br>-Identification of most important needs,<br>-Draft TPP,<br>-Survey to measure stakeholders' preferences,<br>-Revision following feedback,<br>-Round of revisions | Literature (D, analytical accuracy)<br>Expert opinion (D, analytical accuracy),<br>Meeting input (S) | International public organisation (S,C),<br>Representatives of national disease programs (C),<br>Policy-makers (C)<br>Non-profit sector (C)<br>Experts (unspecific) (C),<br>Industry representatives (C),<br>Researchers (C),<br>Clinicians (C) |
| **Ebels, K. et al. (105)** | Journal article | Malaria | / | A combined test | No minimal/optimal test characteristic reported | -Scoping,<br>-Drafting | -Definition problem statement,<br>-Consolidation of findings,<br>-Draft TPP | Interviews with experts (D),<br>Field observations (S,D),<br>Ethnographic interviews (S) | Industry representatives (S,D),<br>Researchers (S,D),<br>Clinicians (S,D)<br>Strategists (S,D),<br>Laboratory experts (S,D),<br>Microbiologists (S,D),<br>Program manager (S,D), |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | | | International public organisation (S,D),<br>Donors (S,D),<br>Representatives of national disease programs (S,D),<br>Policy makers (S,D) |
| **FIND. (107)** | Report | HIV Infection | Bill and Melinda Gates Foundation | A diagnostic test | / | -Scoping,<br>-Drafting<br>-Consensus-building | -Definition problem statement,<br>-Identification of most important needs,<br>-Draft TPP,<br>-Consolidation of findings,<br>-Draft TPP shortened,<br>-Revision following feedback,<br>-Delphi-like approach to gauge stakeholders' agreement with the TPP,<br>-Consensus meeting | Expert opinion (S,D),<br>Meeting input (D) | -Researcher (S,D,C),<br>-Industry representatives (S,C),<br>-International public organisations (S,D,C),<br>-Policy-makers (S,D,C),<br>-Scientific associations (S,D,C) |
| **FIND. (108)** | Report | Human African trypanosomiasis | / | Rapid test for diagnosis and screening | -To define minimal test characteristic as lowest acceptable specification<br>- To define optimal test characteristic as ideal value for that characteristic | -Drafting,<br>-Consensus-building | -Draft TPP,<br>-Delphi-like approach to gauge stakeholders' agreement with the TPP | / | Researchers (C),<br>Non-profit sector (C),<br>Industry representatives (C) |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| **FIND. (109)** | Report | Human African trypanosomiasis | / | A screening test | -To define minimal test characteristic as lowest acceptable specification<br>- To define optimal test characteristic as ideal value for that characteristic | -Drafting,<br>-Consensus-building | -Draft TPP,<br>-Delphi-like approach to gauge stakeholders' agreement with the TPP | / | Researchers (C),<br>International public organisation (C),<br>Non-profit sector (C),<br>Industry representatives (C) |
| **FIND and Forum for Collaborative HIV Research (106)** | Report | Hepatits C | / | 1) HCV nucleic acid amplification diagnostic test<br>2) HCV cAg diagnostic test | -To define minimal test characteristic as lowest acceptable specification<br>- To define optimal test characteristic as ideal value for that characteristic | -Scoping,<br>-Drafting,<br>-Consensus-building | -Priority-setting,<br>-Draft TPP,<br>-Draft TPP shortened,<br>-Delphi-like approach to gauge stakeholders' agreement with the TPP,<br>-Consensus meeting | Literature (D, analytical accuracy),<br>Expert opinion (S,D, accuracy),<br>Available data (analytical accuracy)<br>Models (D, diagnostic accuracy) | Clinicians (C),<br>Implementers (C),<br>Representatives of national disease programs (C),<br>Industry representatives (C),<br>Technical/funding agencies (C),<br>Patient advocates (C),<br>International public organisation (C),<br>Researchers (C) ,<br>Other (C) |
| **Gal, M. et al. (110)** | Journal article | Community-acquired lower respiratory tract infection | Innovative Medicines Initiative IMI-JU-02-2009-04 Infectious Diseases-Diagnostic Tools | POC test | -To define minimal test characteristic as lowest acceptable specification<br>- To define optimal test characteristic as ideal value for that characteristic | -Scoping,<br>-Drafting<br>-Consensus-building | -Priority-setting,<br>-Survey to retrieve input for TPP ,<br>-Meeting with stakeholders,<br>-Draft TPP,<br>-Revision following feedback | Literature (D, diagnostic accuracy),<br>Expert opinion (S,D),<br>Meeting input (D), | Clinicians (S,D,C),<br>Microbiologists (S,C),<br>Industry representatives (S,D,C),<br>Researchers (S,D,C),<br>Market experts (S,D,C), |
| **International Diagnostics Centre (112)** | Published TPP table | HIV | / | A POC test | / | / | / | / | / |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| **International Diagnostics Centre (111)** | Published TPP table | HIV | / | A POC test | / | / | / | / | / |
| **International Diagnostics Centre (113)** | Published TPP table | HIV | / | A POC test | / | / | / | / | / |
| **International Diagnostics Centre (114)** | Published TPP table | HIV and Syphilis | / | A combined test | / | / | / | / | / |
| **Lim, M. et al. (115)** | Journal article | Soil-transmitted helminths | / | 1) A diagnostic test for mapping/monitoring<br>2) A diagnostic test to confirm decision to stop intervention | - To define minimal test characteristics as “must have” requirements<br>- To define optimal characteristics as “nice to have” requirements | -Scoping,<br>-Drafting<br>-Consensus | -Landscaping exercise ,<br>-Definition problem statement,<br>-Draft TPP,<br>Round of revisions,<br>-Meeting with stakeholders,<br>-Revision following feedback | Literature (S,D, diagnostic accuracy)<br>Expert opinion (S,D),<br>Guidelines (D),<br>Meeting input (D) | Researchers (S,D,C),<br>Implementers (S,D,C),<br>Industry representatives (S,D,C) |
| **Nsanzabana, C. et al. (116)** | Journal article | Antimalarial drug resistance | / | Antimalarial drug resistance diagnostic test | -To define minimal test characteristic as lowest acceptable specification<br>- To define optimal test characteristic as ideal value for that characteristic | -Scoping,<br>-Drafting<br>-Consensus-building | -Landscape analysis,<br>-Draft TPP,<br>- Delphi-like approach to gauge stakeholders’ agreement with the TPP,<br>-Consensus-meeting,<br>- Revision following feedback | Literature (S,D, analytical accuracy), | Researchers (C),<br>International public organisation (C),<br>Industry representatives (C) |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| **Pal, S. et al. (117)** | Journal article | Dengue fever | Military Infectious Diseases Research Program, U.S. Army Medical Research and<br>Materiel Command | A diagnostic test | / | -Scoping,<br>-Drafting | -Identification of most important needs,<br>-Draft TPP | Expert opinion (D, diagnostic accuracy),<br>Models (D, diagnostic accuracy),<br>Questionnaire provided to expert (S) | Experts (unspecific) (S,D),<br>Other (S,D) |
| **PATH. (118)** | Report | HIV | / | A self-test | -To define minimal test characteristic as lowest acceptable specification<br>- To define optimal test characteristic as ideal value for that characteristic | -Scoping,<br>-Drafting | -Draft TPP,<br>-Round of revisions | Literature (D, analytical accuracy),<br>Expert opinion (D),<br>Available data (D),<br>Early usability studies (D, analytical accuracy),<br>Laboratory evaluations (D, analytical accuracy) | / |
| **PATH. (119)** | Report | Neglected Tropical Diseases | Bill and Melinda Gates Foundation;<br>UK Department for International Development | 1) Later flow tools<br>2) Nucleic acid amplification tools | -To define minimal test characteristic as lowest acceptable specification<br>- To define optimal test characteristic as ideal value for that characteristic | -Scoping,<br>-Drafting,<br>-Consensus-building | -Definition problem statement,<br>-Draft TPP,<br>-Revision following feedback,<br>-Round of revisions | Literature (S),<br>Interviews with experts (S)<br>Expert opinion (D) | Industry representatives (S,D),<br>Researchers (S,D),<br>Representatives of national disease programs (S,D),<br>Non-profit sector (S,D),<br>Policy makers (S,D),<br>Donors (S,D) |
| **PATH. (120)** | Report | Schistosomiasis | / | A rapid diagnostic test | -To define minimal test characteristic as a must-have | / | / | Literature (D, analytical accuracy),<br>Guidelines (D), | / |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | | | | -To define optimal test characteristic as an ideal value that would make a tool more valuable | | | Expert opinion (D),<br>Field observation (D),<br>Available data (D) | |
| **PATH. (121)** | Report | Schistosomiasis | / | A lateral flow test | -To define minimal test characteristic as a must-have<br>-To define optimal test characteristic as an ideal value that would make a tool more valuable | / | / | Literature (D, analytical accuracy),<br>Expert opinion (D),<br>Field observation (D),<br>Guidelines (D),<br>Available data (D) | / |
| **PATH. (121)** | Report | Schistomiasis | / | A lateral flow test | -To define minimal test characteristic as a must-have<br>-To define optimal test characteristic as an ideal value that would make a tool more valuable | / | / | Literature (D, analytical accuracy)<br>Expert opinion (D),<br>Field observation (D),<br>Guidelines (D),<br>Available data (D) | / |
| **PATH. (123)** | Report | Trachoma | / | A lateral flow rapid diagnostic test | -To define minimal test characteristic as a must-have<br>-To define optimal test characteristic as an ideal value that would make a | / | / | Literature (D, analytical accuracy),<br>Guidelines (D),<br>Available data (D), | / |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | | | | tool more valuable | | | Expert opinion (D) | |
| **PATH. (124)** | Report | Trachoma | / | A lateral flow rapid diagnostic test | -To define minimal test characteristic as a must-have<br>-To define optimal test characteristic as an ideal value that would make a tool more valuable | / | / | Literature (D, analytical accuracy),<br>Guidelines (D),<br>Available data (D),<br>Expert opinion (D) | / |
| **Peck, R. et al. (125)** | Report | Yellow fever | Global Alliance for Vaccines and Immunisation Fund | A field-deployable test | / | -Scoping,<br>-Drafting, | -Definition problem statement,<br>-Draft TPP | Interviews with experts (S,D) | Researchers (S,D),<br>Industry representatives (S,D)<br>Clinicians (S,D),<br>Technical/funding agencies (S,D),<br>Other (S,D) |
| **Porras, A. et al. (126)** | Journal article | Chagas disease | Pan-American Health Organisation ; The Special Programme for Research and Training in Tropical Diseases | 1-2) a POC diagnostic test 3) a treatment-monitoring test | / | -Scoping,<br>-Drafting, | -Meeting with stakeholders,<br>-Definition problem statement,<br>-Draft TPP | Meeting input (S,D)<br>Expert opinion (S) | / |
| **Reipold, E. et al. (127)** | Journal article | Hepatitis C | WHO | A diagnostic test | -To define minimal test characteristic as lowest acceptable specification<br>- To define optimal test characteristic as ideal value | -Scoping,<br>-Drafting,<br>-Consensus-building | -Identification of stakeholders to involve in draft TPP ,<br>-Priority-setting,<br>-Definition TPP domains ,<br>- Draft TPP, | Literature (D, analytical accuracy, diagnostic accuracy),<br>Expert opinion (S,D, diagnostic accuracy), | International public organisation (S,D,C),<br>Researchers (S,D,C),<br>Technical/funding agencies (S,D,C), |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | | | | for that characteristic | | -Delphi-like approach to gauge stakeholders' agreement with the TPP,<br>-Consensus meeting<br>-Survey to measure stakeholders' preferences, | Available data (analytical accuracy)<br>Models (D, diagnostic accuracy) | Industry representatives (S,D,C),<br>Patient advocates (S,D,C),<br>Representatives of national disease programs (S,D,C),<br>Clinicians (S,D,C),<br>Implementers (S,D,C)<br>Policy-makers (C)<br>Program managers (C) |
| **Solomon, A. et al. (128)** | Journal article | Neglected tropical diseases | WHO | 1) A diagnostic tool with mapping and impact monitoring potential<br>2) A diagnostic tool | / | -Drafting | -Draft TPP,<br>- Consensus-meeting | Consensus meeting inputs (D) | Laboratory experts (D,C),<br>Modelers (D,C),<br>Health economists (D,C),<br>Experts (unspecific) (D,C),<br>Program manager (D,C) |
| **Toskin, I. et al. (129)** | Journal article | Sexually transmitted infections | / | 1) A combined diagnostic POC<br>2) POC platforms<br>3) A combined diagnostic test | / | -Scoping,<br>-Drafting,<br>-Consensus-building | -Landscaping exercise ,<br>-Reviewing available literature and data,<br>-Survey to retrieve input for TPP ,<br>-Draft TPP,<br>-Consensus meeting,<br>-Revision following feedback,<br>-Round of revisions | Literature (S,D),<br>Models (D) ,<br>Expert opinion (D), Consensus meeting inputs (D, accuracy),<br>Reports (S),<br>Policies (S),<br>Websites of developers (S),<br>Interviews with experts (S) | Clinicians (D,C),<br>Industry representatives (S,D),<br>International public organisation (D,C),<br>Experts (unspecific) (D,C),<br>Laboratory experts (D,C),<br>Microbiologists (D,C) |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| **UNICEF. (130)** | Report | Pneumonia Acute Respiratory infection | / | A diagnostic aid | / | -Scoping,<br>-Drafting,<br>-Consensus-building | -Defining scope TPP ,<br>-Draft TPP,<br>-Round of revisions,<br>-Survey to measure stakeholders' preferences | Literature (S,D)<br>Expert opinion (D) | Researchers (S,D,C),<br>Industry representatives (S,D,C),<br>International public organisation (S,D,C), |
| **UNICEF. (131)** | Report | E.coli infection | / | Water quality testing product | / | / | / | / | / |
| **UNICEF. (132)** | Report | E.coli infection | / | Water quality testing product | / | / | -Survey to measure stakeholders' preferences | / | / |
| **UNICEF. (133)** | Report | Zika virus | / | A diagnostic test | / | -Scoping,<br>-Drafting | -Defining scope TPP,<br>-Draft TPP,<br>-Round of revisions,<br>-Revision following feedback | Literature (D),<br>Expert opinion (S,D) | Researchers (S,D),<br>Industry representatives (S,D),<br>International public organisation (S,D), |
| **Utzinger, J. et al. (134)** | Journal article | Schistosomiasis | Bill and Melinda Gates Foundation,<br>European Union's Seventh Framework Programme for research, technological development and demonstration | 1) A diagnostic test for mapping-monitoring<br>2) A diagnostic tool | No minimal test characteristic reported | / | / | / | / |
| **WHO (60)** | Report | Tubercolosis | Bill and Melinda Gates Foundation | 1) A diagnostic test<br>2) A diagnostic biomarker test<br>3) A referral screening test<br>4) A drug- | -To define minimal test characteristic as lowest acceptable specification<br>- To define optimal test | -Scoping,<br>-Drafting,<br>-Consensus-building | -Priority-setting,<br>-Draft TPP,<br>-Draft TPP shortened,<br>-Delphi-like approach to gauge | Literature (D, analytical accuracy, diagnostic accuracy), | |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | | | susceptibility test | characteristic as ideal value for that characteristic | | stakeholders' agreement with the TPP,<br>-Consensus meeting | Expert opinion (S,D, analytical accuracy, diagnostic accuracy),<br>Consensus meeting inputs (D),<br>Models (D, analytical accuracy, diagnostic accuracy) | |
| **WHO. (139)** | Report | Meningitis | / | A rapid diagnostic test | / | / | / | / | |
| **WHO and FIND. (137)** | Report | Tubercolosis | / | A test for predicting disease progression | / | -Draft,<br>-Consensus-building | -Meeting with stakeholders ,<br>-Draft TPP,<br>-Survey to retrieve input for TPP ,<br>-Delphi-like approach to gauge stakeholders' agreement with the TPP,<br>-Consensus meeting | Expert opinion (D),<br>Meeting input (D) | |
| **WHO and FIND. (138)** | Report | Buruli ulcer | / | 1) A rapid diagnostic test<br>2) A diagnostic test with treatment-monitoring potential | / | -Scoping,<br>-Draft | -Definition problem statement,<br>-Draft TPP | Expert opinion (S)<br>Meeting input (D) | |
| **WHO (135)** | Report | Taenia solium taeniasis/ cysticercosis | / | 1) A human copro Ag-taeniasis test,<br>2) A combined human Ag/Ab-cysticercosis test<br>3) A porcine Ag-cysticercosis test | / | -Scoping,<br>-Draft | -Definition problem statement,<br>-Draft TPP | Meeting inputs (D),<br>Expert opinion (S) | |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| **WHO et al. (57)** | Report | Severe febrile illness | WHO | A diagnostic platform | / | -Draft,<br>-Consensus-building | -Draft TPP,<br>-Delphi-like approach to gauge stakeholder' agreement with the TPP,<br>-Revision following feedback,<br>-Consensus meeting | Expert opinion (D) | |
| **WHO and FIND. (136)** | Report | Ebola virus | / | A rapid test | / | / | / | / | |

## C.3. Transparency assessment

**Table C-4 Transparency assessment criteria and scores, with TPPs sorted from the most transparent to the least transparent (n=44)**

| | Study mentions type of sources | Study quotes specific literature | Study mentions type of sources for scoping phase | Study describes decision-making steps | Study reports types of stakeholders | Study lists the name of the organizations stakeholders are part of | Study lists the name of each stakeholders involved | Study reports the rationale for inviting certain stakeholders | Study specifies funding body | Total score |
|---|---|---|---|---|---|---|---|---|---|---|
| Nsanzabana, C. et al. (116) | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✗ | 8 |
| WHO (60) | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✗ | ✓ | 8 |
| Dittrich, S. et al. (103) | ✓ | ✓ | ✓ | ✓ | ✓ | ✗ | ✗ | ✓ | ✓ | 7 |
| FIND and Forum for Collaborative HIV Research (106) | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✗ | ✗ | 7 |
| Lim, M. et al. (115) | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✗ | ✗ | 7 |
| Denkinger, C. et al. (99) | ✓ | ✓ | ✓ | ✓ | ✓ | ✗ | ✗ | ✗ | ✓ | 6 |
| Reipold, E. et al. (127) | ✓ | ✓ | ✓ | ✓ | ✓ | ✗ | ✗ | ✗ | ✓ | 6 |
| Toskin, I. et al. (129) | ✓ | ✗ | ✓ | ✓ | ✓ | ✓ | ✓ | ✗ | ✗ | 6 |
| FIND. (107) | ✓ | ✗ | ✓ | ✓ | ✗ | ✓ | ✓ | ✗ | ✓ | 6 |
| WHO and FIND. (137) | ✓ | ✗ | ✗ | ✓ | ✓ | ✓ | ✓ | ✓ | ✗ | 6 |

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| **WHO (135)** | ✓ | ✗ | ✓ | ✓ | ✓ | ✓ | ✓ | ✗ | ✗ | 6 |
| **Peck, R. et al. (125)** | ✓ | ✗ | ✓ | ✓ | ✓ | ✗ | ✗ | ✓ | ✓ | 6 |
| **Denkinger, C. et al. (100)** | ✓ | ✗ | ✓ | ✓ | ✓ | ✗ | ✗ | ✗ | ✓ | 5 |
| **Ding, X. et al. (102)** | ✓ | ✗ | ✓ | ✓ | ✓ | ✗ | ✗ | ✗ | ✓ | 5 |
| **Gal, M. et al. (110)** | ✓ | ✗ | ✓ | ✓ | ✓ | ✗ | ✗ | ✗ | ✓ | 5 |
| **PATH. (119)** | ✓ | ✗ | ✓ | ✓ | ✓ | ✗ | ✗ | ✗ | ✓ | 5 |
| **WHO and FIND. (138)** | ✓ | ✗ | ✓ | ✓ | ✗ | ✓ | ✓ | ✗ | ✗ | 5 |
| **Pal, S. et al. (117)** | ✓ | ✗ | ✓ | ✓ | ✓ | ✗ | ✗ | ✗ | ✓ | 5 |
| **Porras, A. et al. (126)** | ✓ | ✗ | ✗ | ✓ | ✗ | ✓ | ✗ | ✗ | ✓ | 4 |
| **Donadeu, M. et al. (104)** | ✓ | ✗ | ✓ | ✓ | ✓ | ✗ | ✗ | ✗ | ✗ | 4 |
| **PATH. (118)** | ✓ | ✓ | ✗ | ✓ | ✗ | ✗ | ✗ | ✗ | ✓ | 4 |
| **UNICEF. (133)** | ✓ | ✗ | ✓ | ✓ | ✓ | ✗ | ✗ | ✗ | ✗ | 4 |
| **UNICEF. (130)** | ✓ | ✗ | ✓ | ✓ | ✓ | ✗ | ✗ | ✗ | ✗ | 4 |
| **Ebels, K. et al. (105)** | ✓ | ✗ | ✓ | ✓ | ✓ | ✗ | ✗ | ✗ | ✗ | 4 |
| **Chua, A. et al. (98)** | ✓ | ✗ | ✗ | ✓ | ✓ | ✗ | ✗ | ✗ | ✗ | 3 |
| **DIAMETER Project and PATH. (101)** | ✓ | ✓ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✓ | 3 |
| **Solomon, A. et al. (128)** | ✓ | ✗ | ✗ | ✗ | ✓ | ✗ | ✗ | ✗ | ✓ | 3 |
| **FIND. (108)** | ✗ | ✗ | ✗ | ✓ | ✓ | ✗ | ✗ | ✗ | ✗ | 2 |

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| **PATH. (120)** | ✓ | ✓ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | 2 |
| **PATH. (121)** | ✓ | ✓ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | 2 |
| **PATH. (124)** | ✓ | ✓ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | 2 |
| **FIND. (109)** | ✗ | ✗ | ✗ | ✓ | ✓ | ✗ | ✗ | ✗ | ✗ | 2 |
| **PATH. (123** | ✓ | ✓ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | 2 |
| **PATH. (121)** | ✓ | ✓ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | 2 |
| **WHO et al. (57)** | ✓ | ✗ | ✗ | ✓ | ✗ | ✗ | ✗ | ✗ | ✗ | 2 |
| **WHO and FIND. (136)** | ✗ | ✗ | ✗ | ✗ | ✗ | ✓ | ✗ | ✗ | ✗ | 1 |
| **Utzinger, J. et al. (134)** | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✓ | 1 |
| **UNICEF. (131)** | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | 0 |
| **UNICEF. (132)** | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | 0 |
| **WHO. (139)** | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | 0 |
| **International Diagnostics Centre (112)** | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | 0 |
| **International Diagnostics Centre (111)** | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | 0 |
| **International Diagnostics Centre (113)** | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | 0 |
| **International Diagnostics Centre (114)** | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | ✗ | 0 |

# Appendix D

# Test characteristics included in TPPs

**Table D-1 Number and % of test characteristics reported by all TPPs (n=44), sorted by clusters [part 1]**

| **Unmet clinical need** | | **Clinical validity** | |
|---|---|---|---|
| | *n (%)* | | n (%) |
| Intended use | 43 (98) | Diagnostic/testing sensitivity | 31 (70) |
| Medical need | 9 (20) | Diagnostic/testing specificity | 28 (64) |
| Target population | 33 (75) | Positive predictive value | 1 (2) |
| Target user | 33 (75) | Negative predictive value | 1 (2) |
| Medical decision to be influenced | 4 (9) | Field performance | 1 (2) |
| Fit with clinical workflow | 7 (16) | Precision/concordance | 5 (11) |
| Target level of health system | 38 (86) | False recent ratio (%) | 1 (2) |
| Test rationale | 2 (5) | (Generic) sensitivity | 12 (27) |
| Description of test concept | 2 (5) | (Generic) specificity | 12 (27) |
| Proof of concept | 1 (2) | Test performance with disease subgroups | 1 (2) |
| | | What is the risk of an inaccurate test results? | 1 (2) |
| **Analytical performance** | | | |
| Analytical specificity | 18 (41) | Platform throughput | 2 (5) |
| Analytical sensitivity | 11 (25) | Specimen/sample capacity and throughput | 17 (39) |
| Strain specificity | 2 (5) | Manual sample/specimen preparation | 34 (77) |
| Limit of quantification/detection | 13 (30) | Overall sample preparation | 5 (11) |
| Assay throughput | 9 (20) | Need for operator to transfer a precise volume of sample | 7 (16) |
| Volume sample/specimen | 16 (36) | Reagent integration/preparation | 8 (18) |
| Daily throughput (per module) | 3 (7) | Reagent kit (transport, storage and stability, supplies not included in kit) | 3 (7) |
| Assay design/format | 19 (43) | Reagent kit reconstitution/packaging | 8 (18) |
| Reproducibility | 16 (36) | Sample type | 42 (95) |
| Reproducibility near clinical threshold | 1 (2) | Control/comparative reference method | 17 (39) |
| Robustness | 1 (2) | Target molecule to be detected | 17 (39) |
| Interferences | 5 (11) | Type of analysis | 8 (18) |
| Duration of valid sample | 5 (11) | Calibration | 19 (43) |
| Duration of valid result | 7 (16) | Precision/concordance | 6 (14) |
| Result stability | 1 (2) | Quality control | 15 (34) |
| In use stability | 7 (16) | Internal quality control | 16 (36) |
| Quantification/quantitation | 3 (7) | External quality control | 4 (9) |

| | | | |
|---|---|---|---|
| Multiplexing | 7 (16) | Indeterminate test results | 1 (2) |
| Cross reactivity | 1 (2) | Time to test result | 38 (86) |
| Kit quality indicators | 1 (2) | Device failure/invalid rate | 1 (2) |
| (Generic) sensitivity | 12 (27) | Test performance with disease groups | 1 (2) |
| Result | 4 (9) | (Generic) specificity | 12 (27) |
| **Costs** | | **Environmental impact** | |
| Price/cost of individual test | 27 (61) | Environmental footprint | 3 (7) |
| Cost per diagnosis | 2 (5) | | |
| (Capital) cost per instrument | 15 (34) | | |
| | | | |

**Table D-2 Table C-1 Number and % of test characteristics reported by all TPPs (n=44), sorted by clusters [part 2]**

| **Costs** | | **Clinical utility** | |
|---|---|---|---|
| | n (%) | | n (%) |
| Cost of platform to end user | 3 (7) | Intended outcome and linkage to care | 1 (2) |
| Cost of consumables | 3 (7) | | |
| Cost of manufacturing single use device | 3 (7) | | |
| Expected scale of manufacture | 4 (9) | What is the risk of an inaccurate test result? | 1 (2) |
| Potential market | 1 (2) | **Regulatory requirements** | |
| Market segmentation/channels to the market | 10 (23) | Regulatory requirements | 15 (34) |
| Region(s) of commercialisation | 11 (25) | Product registration path | 11 (25) |
| Competitive landscape | 1 (2) | | |
| **Human factors** | | | |
| Assay packaging | 4 (9) | Data analysis | 4 (9) |
| Data capture | 8 (18) | Data export ( connectivity and interoperability, electronics and software) | 18 (41) |
| Data handling | 2 (5) | Data input | 2 (5) |
| Hands-on time | 2 (5) | Instruction for use | 3 (7) |
| Labelling | 3 (7) | Language | 3 (7) |
| Materials used | 3 (7) | Patient identification capability | 5 (11) |
| Rate of errors in device interpretation | 1 (2) | Readout/reading system | 6 (14) |
| Result | 2 (5) | Result documentation-data display | 17 (39) |
| Safety precautions (biosafety requirements) | 12 (27) | Service and support | 10 (23) |
| Supplies needed | 5 (11) | Test outcome (nature) | 21 (48) |
| Tool format and complexity | 12 (27) | Training and education | 33 (75) |
| Unit size | 2 (5) | User interface | 4 (9) |
| User-induced failure rate | 1 (2) | Walkway operation | 2 (5) |

| | | | |
|---|---|---|---|
| Equipment-specific human factors | 7 (16) | Ease of test result interpretation | 10 (23) |
| Service and support | 10 (23) | Test size and weight | 4 (9) |
| Test size and portability | 20 (45) | | |
| **Infrastructural requirements** | | | |
| Ancillary supplies | 7 (16) | Assay packaging | 4 (9) |
| Biosafety requirements | 20 (45) | Clean water | 10 (23) |
| Cold chain | 11 (25) | Environmental tolerance of packaged test kit | 2 (5) |
| External maintenance | 1 (2) | Infrastructural requirements | 6 (14) |
| Instrument-infrastructural requirement | 12 (27) | Maintenance | 15 (34) |
| Additional third-party consumable | 5 (11) | Materials used | 2 (5) |
| Multiuse platform | 2 (5) | Need for additional equipment/test/spare parts | 10 (23) |
| Power requirements | 23 (52) | Reagent kit (transport, storage and stability, supplies not included in kit) | 8 (18) |
| Service and support | 11 (25) | Shipping conditions | 5 (11) |
| Storage conditions and shelf life | 31 (70) | Storage conditions prior utilisation | 1 (2) |
| Thermal tolerance of assay | 3 (7) | Waste disposal | 22 (50) |
| Supplies needed | 5 (11) | Operating conditions | 9 (20) |
| Temperature and humidity | 27 (61) | Stability during transport | 23 (52) |

# Appendix E

# Glossary table

| Term | Description | Source |
|---|---|---|
| **Anaerobic pathogen** | Bacterium that can survive and grow without oxygen. | (344) |
| **Antigen** | Any substance that causes the immune system to produce antibodies against that substance. Common antigens are toxins, chemicals, bacteria, viruses or other substances from the environment. | (345) |
| **Antibody** | Protein being produced by the immune system to destroy the harmful substance (e.g. antigen) as soon as it is detected. | (346) |
| **Bowel flora** | Bacteria and different organisms located in the intestine which help to process food. | (347) |
| **Campylobacter** | Bacterium that can cause food poisoning, intestinal infection diarrhoea and vomiting. | (348) |
| **Celiac disease** | Immune condition in which individuals cannot eat gluten as their immune system would respond by damaging their small intestine. | (349) |
| **Colitis** | Inflammation of the large intestine (i.e. the colon). Also called inflammatory colitis. | (350) |
| **Colonisation** | "The presence of bacteria on a body surface (e.g. skin, mouth, intestines or airway) without causing disease in the person" | (351) |
| **Creatinine** | A chemical waste product that excreted from the body in urine. Abnormal creatinine levels in the blood or urine indicate potential kidney disease. | (352) |
| **Enteral feeding** | A way of providing nutrients to an individual as liquids through a tube placed either in the nose, stomach or small intestine. | (353) |
| ***Escherichia coli*** | Type of bacterium which normally lives in the intestines of healthy individuals. Some types of this bacterium can cause diarrhoea. | (354) |
| **Gastric acid suppressant** | Therapeutic agent (i.e. drug, protein, compound) used in patients with gastrointestinal disorders to reduce the stomach acid production. | (355) |
| **Gastrointestinal neoplasia** | Abnormal mass of tissue in the gastrointestinal tract and other organs of digestion. | (356, 357) |
| **Gram-positive bacterium** | Bacterium that maintains the purple stain during the Gram test. This test helps to identify the type of infection and which antibiotics to administer | (358) |
| **Hypotension** | When blood pressure is lower than normal values. | (359) |
| **Ileus** | Obstruction of the bowel that occurs when there is a stoppage in the digestive tract | (360) |
| **Inflammatory bowel disease** | Term that describes two conditions (i.e. ulcerative colitis and Chron's disease) where there is inflammation (e.g. swelling and irritation) of the gut. | (361) |
| **Irritable bowel disease** | Health condition where the nutrients move either too fast or too slow through the intestines causing abdominal pain and change in bowel habits. | (362) |
| **In vitro** | Experiments conducted with biological components (e.g. cells, tissues) that have been removed from a living organism. | (363) |
| **In vivo** | Studies conducted on living organisms. | (363) |
| **Multiplex test** | A test detecting multiple targets simultaneously | (364) |
| **Overflow** | Inability to control urination. | (365) |
| **Oxidative stress** | Imbalance between free radicals and antioxidants in a body. | (366) |

| | | |
|---|---|---|
| **Salmonella** | Bacterium that causes foodborne illness with symptoms ranging from fever to diarrhoea. | (367) |
| **Shigella** | Bacterium that causes infection of the intestine linings. | (368) |
| **Spore** | Cell produced by certain bacteria, fungi and plants as defence mechanism. | (369) |
| **Strain** | Variant or subtype of an organism. | (370) |
| **Toxin megacolon** | “A nonobstructive dilation of the colon, which can be total or segmental and is usually associated with systemic toxicity” | (371) |
| **White blood cell count** | It measures the number of white cells in the blood. | (372) |

# Appendix F

# Online survey "Identifying unmet clinical needs in the diagnosis of Clostridioides difficile infection"

**Information Sheet**

You are being invited to participate in a questionnaire which aims to identify any unmet clinical needs in the diagnosis of Clostridioides difficile (*C. difficile*) infection in the UK and to map the current clinical practice. Please read the following information carefully.

**Introduction**. This questionnaire is being conducted as part of a PhD by Paola Cocco at the University of Leeds, under the supervision of Dr Bethany Shinkins, Dr Michael Messenger, Dr Kerrie Davies, Professor Robert West and Dr Alison Smith. This questionnaire was also designed with feedback from Professor Mark Wilcox.

**Why have I been chosen?** We would like to hear your views as you have experience of *C. difficile* infection (medical doctors, nurses, GPs and biomedical scientists) at the hospital and community care setting.

**What will I have to do if I take part?** If you decide to take part, you will be asked your consent to participate at the beginning and again at the end of the questionnaire. This questionnaire will take you approx. 5-10 minutes to complete and does not require you to search for any additional information. Once you submit, it will not be possible to withdraw your responses.

This questionnaire will ask you some questions on care practice for *C. difficile* infection in the clinical setting where you work, your opinion on current diagnostics and if there are any problems with them, what impact these problems have on patient health and how to better address these clinical needs.

**Do I have to take part?** No. Your participation is entirely voluntary and you can withdraw at any point up to submission.

**What are the risks of taking part?** We believe there are no known risks associated with this questionnaire; however, as with any online related activity the risk of a breach is always possible.

**How will my answers be used**? Your participation will remain confidential, and only anonymised data will be published. Further information is available via the University of Leeds Privacy Notice. . Main analyses will be conducted by Paola Cocco with the support of the research team.

**Who has reviewed the study?** Ethical approval has been sought from the School of Medicine Research Ethics Committee at the University of Leeds (reference number MREC 19-047).

**Please continue to complete the survey**.

Many thanks!

**Q1.** Do you consent to participate to this questionnaire?*[35]

- ☐ Yes
- ☐ No

*[If answer to Q1, move to Final Page]*

[35] * indicates a mandatory question

**Respondent Details**

**Q2.** What is your main job title?*

a) General Practitioner
b) Consultant
c) Trainee/Staff Doctor
d) Infection Disease Doctor
e) Medical Microbiologist
f) Nurse
g) Infection Control Nurse
h) Biomedical Scientist
i) Healthcare Scientist
j) Clinical Scientist
k) Other (please specify)

*[If answer to Q2 is a), move to Q3] [If answer to Q2 is b-k), move to Q2a]*

**Q2a.** Please select your specialty*

□ Cardiology
□ Chemical Pathology
□ Endocrinology
□ Gastroenterology
□ Geriatric (Elderly Medicine)
□ General (Internal Medicine)
□ Haematology
□ Infectious Diseases
□ Infection Prevention and Control
□ Intensive Care
□ Medical Microbiology and Virology
□ Neurology
□ Oncology
□ Renal Medicine/Nephrology
□ Respiratory Medicine
□ Rheumatology
□ Urology
□ Other (please specify)

**Q3.** How many years of experience do you have working in your clinical section?*

□ 0 to 4 years
□ 5 to 10 years
□ 11 to 15 years
□ More than 15 years

**Q4.** Where do you work?*

□ England
□ Scotland
□ Northern Ireland
□ Wales

**Q5.** In which clinical setting do you work?*

a) Hospital
b) GP Medical Practice
c) Sheltered Accommodation
d) Residential Home
e) Care Services at Home
f) Nursing Home
g) Hospital-based laboratory
h) Independent laboratory
i) Other (please specify)

*[If answer to Q2 is a, move to Q11] [If answer to Q2 is b-d, move to Q6] [If answer to Q2 is f-g, move to Q18]*

*[If answer to Q2 is e, h-j, move to Q7] [If answer to Q2 is k, move to Q6]*

**Field of Expertise**

**Q6**. Please select which area of expertise you consider yourself more familiar with.*

a) Laboratory practice
b) Clinical practice

*[If answer is a), move to Q7] [If answer is b), move to Q18]*

**Laboratory-based clinicians**

**Section 1: Description of Laboratory Practice: Testing Option(s)**

We would like to gather your views about laboratory practice in diagnosing Clostridioides difficile (*C. difficile*) infection in the laboratory where you work.

**Q7.** Which of the following diagnostic tests are used in your laboratory for individuals with suspected *C. difficile* infection? Please select one or more options.

□ Toxigenic culture
□ Culture
□ Cytotoxicity assay
□ Enzyme immunoassay (EIA) for toxin A/B
□ Enzyme immunoassay (EIA) FOR Glutamate Dehydrogenase
□ Nucleic Acid Amplification Test (NAAT)
□ Other (please specify)
□ I do not know

**Q8.** Would you run a combination of diagnostic tests ('testing algorithm') for a patient with suspected *C. difficile* infection?

a) Yes
b) No
c) I do not know

[if a), move to Q9, if b-c) move to Q27]

**Section 1: Description of Laboratory Practice: Testing Algorithm**

**Q9.** For each step of the testing algorithm, please describe in the best way you can what would you do if the test result was positive or negative. If your laboratory uses more testing algorithms please describe them separately.

_____
_____

**Q10.** Do you know why this particular testing algorithm is used in your laboratory? If so, please explain.

_____
_____

[Move to Q30]

**Ward-based clinicians**

**Section 1: Description of Current Practice**

We would like to gather your views about current practice in diagnosing Clostridioides difficile (*C. difficile*) infection in the clinical setting where you work.

**Q18.** Setting aside treatment regimens, what **symptoms** would lead you to suspect that an individual in your clinical setting has *C. difficile* infection? Please select one or more options.

- □ Watery stools
- □ Fever
- □ Abdominal cramps
- □ Leucocytosis
- □ Hypoalbuminemia
- □ Unexplained diarrhoea
- □ Other (please specify)
- □ I do not know

**Q19.** In which age group do you most often encounter individuals with suspected *C. difficile* infection? Please select one or more options.

- □ Young children (< 2 years old)
- □ Children (2 to 18 years old)
- □ Adults (18 to 65 years old)
- □ Older adults (>65 years old)
- □ I do not know

**Q20.** As part of this process, would you request any diagnostic test?

a) Yes
b) No

[If a), move to Q20a] [if b), move to Q25]

**Q20a**. Would you presumptively isolate a patient suspected with *C. difficile* infection while awaiting test results?

a) Yes
b) No
c) I do not know

[Move to Q21]

**Description of Current Practice: Diagnostic Pathway (part 1)**

**Q21.** Do you know which test(s) the laboratory uses to diagnose individuals suspected with *C. difficile* infection?

a) Yes
b) No

[If a), move to Q22] [if b), move to Q26]

**Description of Current Practice: Diagnostic Pathway (part 2)**

**Q22.** Which of the following diagnostic tests would you request to test a patient with suspected *C. difficile* infection? Please select one or more options.

- □ Toxigenic culture
- □ Culture
- □ Cytotoxicity assay
- □ Enzyme immunoassay (EIA) for toxin A/B
- □ Enzyme immunoassay (EIA) FOR Glutamate Dehydrogenase
- □ Nucleic Acid Amplification Test (NAAT)
- □ Other (please specify)

□ I do not know

**Q23.** Would you request a combination of diagnostic tests ('testing algorithm') for a patient with suspected *C. difficile* infection?

d) Yes
e) No
f) I do not know

[if a), move to Q24, if b-c) move to Q26]

**Q24.** For each step of the testing algorithm, please describe in the best way you can what would you do if the test result was positive or negative. If your laboratory uses more testing algorithms please describe them separately.

_____
_____

[Move to Q26]

**Q25.** For each of the following scenarios, what would you consider to diagnose a patient suspected with *C. difficile* infection (e.g. patient demographics, risk factors, symptoms, underlying conditions and treatment)?

| | |
|---|---|
| CDI is likely to be present | |
| *C. difficile* could be present | |
| CDI is very UNLIKELY to be present | |

*Cloistridioides difficile infection = CDI*

*Terminology for each diagnostic scenario is based on Department of Health and Social Care document "Update guidance on the diagnosis and reporting of Clostridium Difficile"(https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/215135/dh_133016.pdf)*

*[Move to Q26]*

**Description of Current Practice: Diagnostic Pathway (part 3b)**

**Q26.** Using a scale from 1= "Not at all confident" to 5="Very confident", how confident do you usually feel about the test results for *C. difficile* infection you receive?

□ Not at all confident 1
□ Not very confident 2
□ Neither 3
□ Fairly confident 4
□ Very confident 5
□ I do not know

**Q27.** What is your confidence in test results for *C. difficile* infection based on? Please explain.

_____
_____

**Q28.** What would make you more confident about test results for *C. difficile* infection?

_____
_____

*[Move to Q29]*

**Description of Current Practice: Diagnostic Pathway (part 4)**

**Q29.** What would you consider to initiate de-escalation of infection control measures for a patient confirmed with *C. difficile* infection (e.g. test results, patient demographics, risk factors, symptoms, underlying conditions and treatment)?

_____
_____

*[Move to Q30]*

**General Practitioners**

**Section 1: Description of Clinical Practice (GP)**

We would like to gather your views about current practice in diagnosing Clostridioides difficile *(C. difficile)* infection in the clinical setting where you work.

**Q11.** What symptoms would lead you to suspect that an individual in your clinical setting has *C. difficile* infection? Please select one or more options.

- □ Watery stools
- □ Fever
- □ Abdominal cramps
- □ Leucocytosis
- □ Hypoalbuminemia
- □ Unexplained diarrhoea
- □ Other (please specify)
- □ I do not know

**Q12.** In which age group do you most often encounter individuals with suspected *C. difficile* infection? Please select one or more options.

- □ Young children (< 2 years old)
- □ Children (2 to 18 years old)
- □ Adults (18 to 65 years old)
- □ Older adults (>65 years old)
- □ I do not know

**Q13.** As part of this process, would you request any diagnostic test?

a) Yes
b) No
c) I do not know

*[If a), move to Q13a] [If b-c), move to Q14]*

**Q13a.** Would you suggest a patient suspected with *C. difficile* infection to follow hand hygiene measures after requesting a diagnostic test?

a) Yes
b) No
c) I do not know

*[Move to Q15]*

**Description of Clinical Practice: Diagnostic Pathway (part a)**

**Q14.** For each of the following scenarios, what would you consider to diagnose a patient suspected with *C. difficile* infection (e.g. patient demographics, risk factors, symptoms, underlying conditions and treatment)?

*Cloistridioides difficile infection = CDI*

*Terminology for each case is based on Department of Health and Social Care document "Update guidance on the diagnosis and reporting of Clostridium Difficile"(https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/215135/dh_133016.pdf)*

| | |
|---|---|
| CDI is likely to be present | |
| C. difficile could be present | |
| CDI is very unlikely to be present | |

*[Move to Q15]*

**Description of Clinical Practice: Diagnostic Pathway (part b)**

**Q15.** Using a scale from 1= "Not at all confident" to 5="Very confident", how confident do you usually feel about the test results for *C. difficile* infection you receive?

- ☐ Not at all confident 1
- ☐ Not very confident 2
- ☐ Neither 3
- ☐ Fairly confident 4
- ☐ Very confident 5
- ☐ I do not know

**Q16.** What is your confidence in test results for *C. difficile* infection based on? Please explain.

_____
_____

**Q17.** What would make you more confident about test results for *C. difficile* infection?

_____
_____

*[Move to Q30]*

**Common sections**

**Section 2: Problems with Diagnosing Clostridioides difficile infection**

In this section we are interested in identifying any problems or difficulties with diagnosing *C. difficile* infection in the clinical setting where you work.

**Q30**. In your view, are there any problems in diagnosing C. *difficile* infection? If so, please describe them here.

_____
_____

**Q31**. In your view, are there any problems with current diagnostic tests for C. *difficile* infection?*

a) Yes
b) No
c) I do not know

*[If answer is a) move to Q32] [If answer is b-c) move to Q34]*

**Section 2. Problems with Current Diagnostic Tests**

**Q32.** In relation to the diagnostic test(s) that is available to you in your clinical settings, what are the limitations? Please select one or more options.*

- □ Long turnaround time
- □ Too expensive
- □ Poor ability to confirm if a patient has *C. difficile* infection
- □ Poor ability to confirm if a patient does not have *C. difficile* infection
- □ Difficult to interpret test results
- □ Other (please specify)
- □ I do not know

**Q33.** In relation to the diagnostic tests that are available to you in your laboratory, what problems do these limitations cause? Please select one or more options.*

- □ Longer length of (hospital) stay
- □ Inappropriate antibiotic prescription
- □ Delays in administering treatment to patients
- □ Differences in case reporting across laboratories
- □ Potential spread of infection
- □ Other (please specify)
- □ I do not know

*[Move to Q34]*

**Section 3: Room for Improvement in Current Diagnostic Tests?**

**Q34**. Could a better diagnostic test solve any of the current issues with diagnosing *C. difficile* infection?*

a) Yes
b) No
c) I do not know

*[If answer is a) move to Q36] [If answer is b) move to Q35a]*

_____
_____

**Q35a**. Please explain your answer.

*[If answer to Q35 is b), move to Q38]*

**Section 4: Description of Ideal Diagnostic Test**

**Q36**. How would a better diagnostic test solve current issues with diagnosing C. *difficile* infection?

_____
_____

We would now like to collect your thoughts on the ideal features of a new diagnostic test for C. *difficile* infection.

**Q37**. Hypothetically, what would be the ideal characteristics of a new diagnostic test for C. *difficile* infection in relation **to your clinical setting**?

| | |
|---|---|
| Ideal type of sample (e.g. stool) | |
| Ideal technique for obtaining the sample (e.g. swab) | |
| Ideal technique for transporting the sample | |
| Ideal turnaround time | |
| Ideal positioning of a new test into care pathway | |
| Acceptable cost of a new test (£) | |
| Other ideal features of a new test | |

**Consent to Submit**

**Q38.** Do you consent to submit this questionnaire? Once you submit the questionnaire it will not be possible to withdraw your responses.*

a) Yes
b) No

*[if answer is a), move to Final page] [if answer is b), move to 'End of the Questionnaire']*

**End of the Questionnaire**

Many thanks for filling in the questionnaire. Your answers will be deleted as you have chosen not to submit the questionnaire.

Please close this webpage without pressing the button 'Finish'.

If you have any queries, please contact via email Paola Cocco (umpc@leeds.ac.uk)

*[END QUESTIONNAIRE WITHOUT SUBMITTING IT]*

**Final page**

Many thanks for filling in the questionnaire. Your answers will give us great insights on current UK practice for *C. difficile* infected patients.

If you know someone who might be interested in taking part in this questionnaire, please share this link https://leeds.onlinesurveys.ac.uk/c-difficile-diagnostics.

If you have any queries, please contact via email Paola Cocco (umpc@leeds.ac.uk)

# Appendix G

# Ethical approval – supporting information

**From:** Rachel De Souza [Medicine] on behalf of Medicine and Health Univ Ethics Review
**Sent:** 09 October 2019 15:23
**To:** Paola Cocco
**Cc:** Medicine and Health Univ Ethics Review
**Subject:** RE: Ethical approval-advice
**Attachments:** HRA decision tool results.pdf; SoMREC Submission

Checklist

Hi Paola

I can see that the HRA decision tool "Do I need NHS REC approval?" outcome is that you do not require NRES (NHS) ethics approval however, to ascertain if you study is research *by NHS standards* can you please also complete the other HRA decision tool at http://www.hra-decisiontools.org.uk/research/ .

If the outcome of that tool is that your study is not considered to be research by NHS standards, then please follow the attached SoMREC submission checklist.
Many thanks
Rachel

~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~

**Rachel de Souza, Research Ethics & Governance Administrator,** The Secretariat, Room 9.29, Level 9, Worsley Building, Clarendon Way, University of Leeds, LS2 9NL, Tel: 0113 3431642, r.e.desouza@leeds.ac.uk

---

**From:** Paola Cocco
**Sent:** 09 October 2019 14:33
**To:** Medicine and Health Univ Ethics Review <FMHUniEthics@leeds.ac.uk>
**Subject:**Ethics

Hello

I am Paola Cocco, a PhD student in health economics. I am writing you to ask an advice on seeking approval from either University of Leeds or from the NHS Health Authority. As part of my project, I will send a qualitative online survey to NHS staff (e.g. clinicians, nurses, laboratory experts) and patient representatives (PPI) in order to elicit an unmet clinical need within the field of Clostridium Difficile Infection diagnostic tests. Specifically, I plan to ask NHS staff which patient population they think is currently affected by poor diagnosis, whether there is an existing diagnostic test which could solve this issue and what are the characteristics of an ideal test which will help addressing this issue. NHS staff will be recruited by virtue of their professional role using snowball sampling.
I have checked the HRA online decision tool and it does not seem that I would need a NHS ethical approval (please see file attached). Could you please let me know whether I will need an approval from University of Leeds or from NHS Health Authority?

Thanks for your help.
Sincerely,
~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~

MRC Medical Research Council

NHS Health Research Authority

## Is my study research?

**To print your result with title and IRAS Project ID please enter your details below:**

Title of your research:

Developing a framework based on early economic evaluation to inform target product profiles for new diagnostic tests

IRAS Project ID (if available):

You selected:

- **'No'** - Are the participants in your study randomised to different groups?
- **'No'** - Does your study protocol demand changing treatment/ patient care from accepted standards for any of the patients involved?
- **'No'** - Are your findings going to be generalisable?

**Your study would NOT be considered Research by the NHS.**

You may still need other approvals.

Researchers requiring further advice (e.g. those not confident with the outcome of this tool) should contact their R&D office or sponsor in the first instance, or the HRA to discuss your study. If contacting the HRA for advice, do this by sending an outline of the project (maximum one page), summarising its purpose, methodology, type of participant and planned location as well as a copy of this results page and a summary of the aspects of the decision(s) that you need further advice on to the HRA Queries Line at HRA.Queries@nhs.net.

For more information please visit the Defining Research table.

Follow this link to start again.

Print This Page

NOTE: If using Internet Explorer please use browser print function.

# Appendix H

# Ethical approval from School of Medicine Research Ethics Committee

**From:** Rachel De Souza [Medicine] on behalf of Medicine and Health Univ Ethics Review
**Sent:** 04 March 2020 15:34
**To:** Paola Cocco
**Cc:** Medicine and Health Univ Ethics Review; Bethany Shinkins
**Subject:** RE: MREC 19-047 Study Approval

**Importance:** High

Dear Paola

**MREC 19-047 - Identifying unmet clinical needs for a test for Clostridium Difficile Infection**

I am pleased to inform you that the above research ethics application has been reviewed by the School of Medicine Research Ethics Committee (SoMREC) Committee and on behalf of the Chairs, I can confirm a favourable ethical opinion based on the documentation received at date of this email.

***Please retain this email as evidence of approval in your study file.***

Please notify the committee if you intend to make any amendments to the original research as submitted and approved to date. This includes recruitment methodology; all changes must receive ethical approval prior to implementation. Please see https://leeds365.sharepoint.com/sites/ResearchandInnovationService/SitePages/Amendments.aspx or contact the Research Ethics Administrator for further information (FMHUniEthics@leeds.ac.uk ) if required.

Ethics approval does not infer you have the right of access to any member of staff or student or documents and the premises of the University of Leeds. Nor does it imply any right of access to the premises of any other organisation, including clinical areas. The committee takes no responsibility for you gaining access to staff, students and/or premises prior to, during or following your research activities.

*Please note:* You are expected to keep a record of all your approved documentation, as well as documents such as sample consent forms, risk assessments and other documents relating to the study. This should be kept in your study file, which should be readily available for audit purposes. You will be given a two week notice period if your project is to be audited.

It is our policy to remind everyone that it is your responsibility to comply with Health and Safety, Data Protection and any other legal and/or professional guidelines there may be.

I hope the study goes well.

Best wishes

Rachel

***On behalf of Dr Naomi Quinton and Dr Anthony Howard, co-Chairs, SoMREC***

~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~

**Rachel de Souza, Lead Research Ethics & Governance Administrator,** The Secretariat, Room 9.29, Level 9, Worsley Building, Clarendon Way, University of Leeds, LS2 9NL, Tel: 0113 3431642, r.e.desouza@leeds.ac.uk
~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~

# Appendix I

## Additional results - online survey

This appendix provides additional results from the online survey presented in Chapter 4.

**Table I-1 Summary statistics (n, %) of questions related 'Description of Clinical Practice' section, sorted by respondent subgroups**

| Would you request a diagnostic test upon suspicion of CDI? | | | | |
|---|---|---|---|---|
| | **Total responses (n=37)** | **Consultants and Doctors (n=9)** | **Nurses (n=23)** | **GPs and others (n=5)** |
| **Yes** | 35 (95) | 9 (100) | 22 (96) | 4 (8) |
| **No** | 0 (0) | 0 (0) | 0 (0) | 0 (0) |
| **I do not know** | 2 (5) | 0 (0) | 1 (4) | 1 (2) |
| **Would you pre-emptively start infection-control measures** | | | | |
| Yes | 36 (97) | 9 (100) | 22 (96) | 5 (100) |
| No | 0 (0) | 0 (0) | 0 (0) | 0 (0) |
| I do not know | 1 (3) | 0 (0) | 1 (4) | 0 (0) |
| **Do you know which test(s) the laboratory uses to diagnose CDI?** | | | | |
| | **Total responses (n=33)** | **Consultants and Doctors (n=9)** | **Nurses (n=23)** | **Others (n=1)** |
| **Yes** | 29 (88) | 9 (100) | 20 (87) | 0 (0) |
| **No** | 4 (12) | 0 (0) | 3 (13) | 1 (100) |
| **I do not know** | 0 (0) | 0 (0) | 0 (0) | 0 (0) |

**Figure I-1 Availability of diagnostic tests for CDI in absolute number (n) and as a proportion of respondents selecting that option within each subgroup (%).**

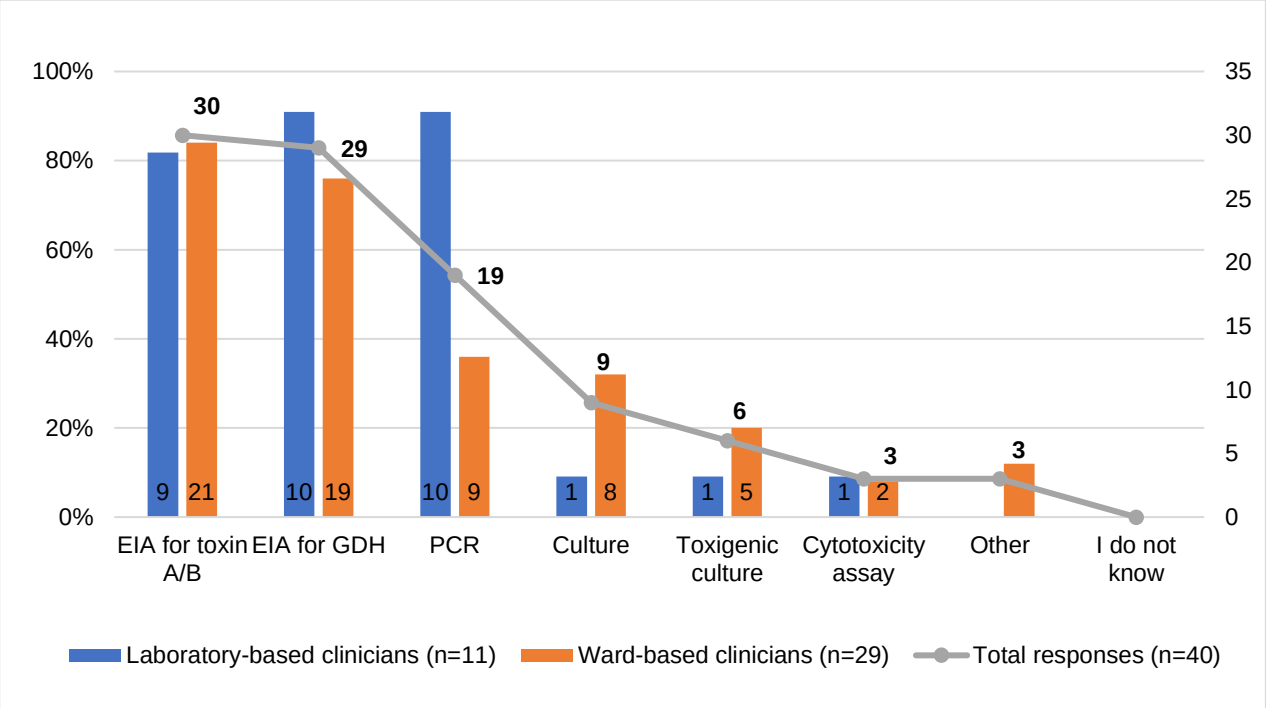



**Table I-2 Summary statistics (n, %) of a question related to practice of running testing algorithms for CDI, sorted by respondent subgroups**

| | Would you run a combination of diagnostic tests for CDI | | |
|---|---|---|---|
| | **Total responses (n=40)** | **Laboratory-based clinicians (n=11)** | **Ward-based clinicians (n=29)** |
| **Yes** | 31 (78) | 10 (91) | 21 (72) |
| **No** | 3 (8) | 1 (9) | 2 (7) |
| **I do not know** | 6 (15) | 0 (0) | 6 (21) |

**Table I-3 Summary statistics (n, %) of question related to *'Problems with diagnostic tests for CDI'* section, sorted by respondent subgroups**

| Are there any problems with diagnostic tests for CDI? | | | | |
|---|---|---|---|---|
| | **Total responses (n=48)** | **Laboratory-based clinicians (n=11)** | **Ward-based clinicians (n=33)** | **GPs (n=4)** |
| **Yes** | 19 (40) | 6 (55) | 11 (33) | 2 (50) |
| **No** | 12 (25) | 5 (45) | 6 (19) | 1 (25) |
| **I do not know** | 17 (35) | 0 (0) | 16 (48) | 1 (25) |
| **Is there room for improvement for new diagnostic tests for CDI?** | | | | |
| **Yes** | 19 (40) | 7 (64) | 10 (30) | 2 (50) |
| **No** | 9 (19) | 4 (36) | 4 (12) | 1 (25) |
| **I do not know** | 20 (42) | 0 (0) | 19 (58) | 1 (25) |

**Table I-4 Desirable specifications new diagnostic tests for CDI should ideally possess according to respondents (n=19), sorted by themes and subthemes in absolute number (n) and as a proportion of respondents which discussed that theme (%).**

| **Ideal type of sample** | **n (%)** |
|---|---|
| Stool | 15 (79) |
| Blood | 4 (21) |
| Saliva | 1 (5) |
| Skin microbes | 1 (5) |
| Other | 1 (5) |
| N/A | 1 (5) |
| **Ideal technique for obtaining the sample** | |
| Rectal swab | 7 (37) |
| Stool collection | 5 (26) |
| Others | 5 (26) |
| N/A | 2 (11) |
| **Ideal technique for transporting the sample** | |
| Tube | 3 (16) |
| Chute | 1 (5) |
| Porter | 1 (5) |
| Ward-based/ no need for transportation | 2 (11) |
| Ambient/unrefrigerated | 2 (11) |
| Other | 6 (32) |
| N/A | 2 (11) |
| **Ideal turnaround time** | |
| < 1 hour | 8 (42) |
| < 2 hours | 2 (9) |

| | |
|---|---|
| 2 – 4 hours | 2 (11) |
| 6 hours | 2 (11) |
| 24 hours | 3 (16) |
| N/A | 1 (5) |
| **Ideal positioning of a new test into care pathway** | |
| Point-of-care | 8 (42) |
| Symptom onset | 4 (21) |
| N/A | 6 (32) |
| **Acceptable cost of a test** | |
| ≤ £10 | 7 (37) |
| < £30 | 3 (16) |
| <50£ | 1 (5) |
| £80 | 1 (5) |
| Free | 2 (11) |
| N/A | 4 (21) |

The absolute number of respondents might not add up to 19 in case respondents proposed more than one subtheme for each desirable specification

# Appendix J

# Search strategy – rapid literature review decision models for CDI diagnostics

This appendix provides the search strategy of the rapid literature review of decision models for diagnostic tests for CDI presented in Chapter 5.

**Table J-1 Results from searches conducted on 18th February 2021 on MEDLINE (Ovid) and EMBASE**

| # | Searches | Results |
|---|---|---|
| 1 | Clostridium difficile/ | 10245 |
| 2 | Clostridium Infections/ | 9573 |
| 3 | C#diff*.ti,ab,kw | 459 |
| 4 | diarrhea/ | 49798 |
| 5 | exp models,economic/ | 15859 |
| 6 | markov chains/ | 15309 |
| 7 | Markov.ti,ab,kf | 53771 |
| 8 | Monte carlo method/ | 30285 |
| 9 | Monte carlo.ti,ab,kf | 53771 |
| 10 | Exp Decision Theory/ | 12633 |
| 11 | Simulation*.ti,ab,kf | 392165 |
| 12 | Economic model*.ti,ab,tw | 3576 |
| 13 | Decision model*.ti,ab,tw | 2662 |
| 14 | (decision* ADJ2 (tree* OR analy* OR model*)).ti,ab,kw | 79110 |
| 15 | (cost* ADJ1 (benefit* OR utility* OR consequence* OR effectiveness)).ti,ab,kw | 79110 |
| 16 | OR/5-15 | 544368 |
| 17 | 1 AND 16 | 116 |
| 18 | 2 AND 16 | 105 |
| 19 | 3 AND 16 | 331 |
| 20 | 4 AND 16 | 17 |
| 21 | OR/17-20 | 471 |
| 22 | Limit 21 to English language | 453 |

# Appendix K

# Results – rapid literature review decision models for CDI diagnostics

This appendix provides the results from the rapid literature review of decision models for diagnostic tests for CDI presented in Chapter 5, including: (i) results search strategy; (ii) details of each of the decision models for diagnostic tests for CDI included in the rapid literature review; and (iii) scores for quality of methodological reporting per item of the CHEERS checklist for each of the included studies (167).

**Figure K-1 PRISMA flow diagram of the rapid literature review search results**

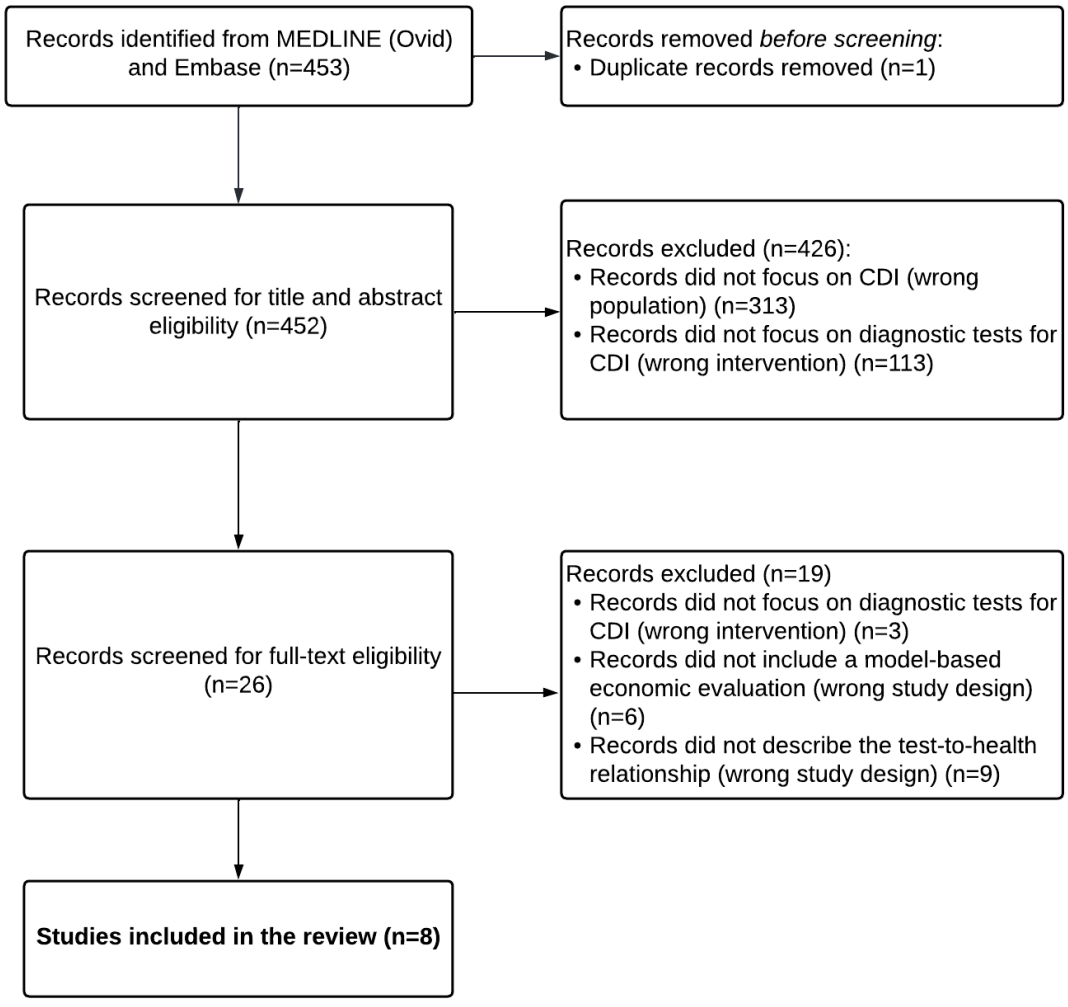


**Table K-1 Summary table of the studies included in the literature review of decision models for CDI diagnostic tests**

| First author, year and country | Population | Intervention(s) | Comparator(s) | Model type | Perspective and time horizon | Analysis and outcomes |
|---|---|---|---|---|---|---|
| **Bartsch, S.M. et al. (264)** | Hypothetical cohort of 1,000 patients admitted to hospital | Screening with PCR | No screening | Decision tree | – Hospital and third payer perspective<br>– Time horizon not specified | – CUA, CEA<br>– QALY, cost per case averted (i.e. the difference in cost divided by the number of cases that screening would prevent) |
| **Schroeder (2013)(267), USA** | Hypothetical cohort of 10,000 adult inpatients suspected of having CDI | – PCR<br>– GDH EIA followed by PCR (if GDH EIA positive)<br>– GDH EIA followed EIA toxin A/B | – Batch PCR,<br>– EIA toxin A/B,<br>– Direct tissue culture cytotoxicity<br>Two treatment strategies: treat none, treat all | Decision tree | – Healthcare perspective<br>– "In the timespan of weeks" time-horizon | – CEA, CBA<br>– Cost per true case diagnosed and treated, net cost of diagnosis and treatment (including cost of missed cases) |
| **Bartsch (2015)(235), USA** | Hypothetical cohort of 100 inpatients who had submitted samples for testing | – EIA toxin A/B<br>– GDH followed by EIA toxin A/B<br>– PCR<br>– GDH EIA followed by EIA toxin A/B, followed by PCR for confirmation of indeterminate results | None specified | Decision tree | – Healthcare perspective<br>– Time horizon not specified | – CEA<br>– Number of (i) timely, (ii) delayed, (iii) unnecessary treatments for CDI; unnecessary additional bed days; opportunity cost for lost bed days |
| **Saab (2015)(269), USA** | Hypothetical hospital inpatients with cirrhosis | Screening all patients with EIA toxin A/B regardless of symptoms | Screen only symptomatic patients with EIA toxin A/B | Not specified | – Perspective not specified<br>– Time horizon not specified | – Type of analysis not reported<br>– Outcome evaluated not reported |

| | | | | | | |
|---|---|---|---|---|---|---|
| **Freeman (2017)(265), UK** | Hypothetical cohort of adult inpatients with suspected gastroenteritis | GPP | Conventional testing | Decision tree | – Healthcare perspective<br>– Two weeks' time-horizon | – CUA<br>– QALY loss, total costs |
| **Schechner (2017)(270), Israel** | Hospitalised patients in the internal medicine department (n=350) | Four total strategies<br>– Two testing methods (e.g. GDH EIA followed by EIA toxin A/B; PCR)<br>– Two infection control measures (contact isolation in multi-bed rooms or single bed isolation/cohorting) | GDH EIA followed by EIA toxin A/B plus contact isolation in multiple-bed rooms | Stochastic dynamic model | – Healthcare perspective<br>– One year | – CEA<br>– Average daily number of infected patients, total costs |
| **Jones (2020)(266), UK** | Hypothetical cohort of 100 inpatients with infectious diarrhoea | PCR point-of-care followed by EIA toxin A/B | GDH EIA followed by PCR test and EIA toxin A/B | Decision tree | – Healthcare perspective<br>– Two weeks' time-horizon | – Cost-consequence analysis<br>– Time-to-result, time in single room isolation, total length of hospital stay, cost of bed days, costs of diagnostic testing, total costs |
| **Xuan (2020)(268), USA** | 65 year-old adult with diarrhoea, tested in an community setting | – Standalone PCR<br>– GDH + EIA toxin A/B<br>– NAAT + EIA toxin A/B<br>– GDH + NAAT | None specified | Decision tree | – Societal perspective<br>– 19 years' time-horizon | – CUA<br>– QALYs, total CDI cases, total costs |

CBA – cost-benefit analysis; CEA – cost-effectiveness analysis; CUA – cost-utility analysis; EIA – enzyme immunoassay; GDH – glutamate dehydrogenase; GPP – gastrointe
chain reaction; QALY – quality-adjusted life year

**Table K-2 CHEERS checklist scores for each of the included studies in the rapid literature review of decision models for diagnostic tests for CDI (n=8)(167)**

| | Study (publication year) | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| | Bartsch (2012)(264) | Schroeder (2013)(267) | Bartsch (2015)(235) | Saab (2015)(269) | Freeman (2017)(265) | Schechner (2017)(270) | Jones (2020)(266) | Xuan (2020)(268) |
| **1. Title** | 0 | 1 | 0 | 0 | 0.5 | 1 | 1 | 1 |
| **2. Abstract** | 0.5 | 0.5 | 1 | 0.5 | 1 | 1 | 1 | 1 |
| **3. Background and objectives** | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 1 |
| **4. Target population and subgroups** | 0.5 | 1 | 1 | 0.5 | 1 | 1 | 1 | 1 |
| **5. Settings and location** | 0 | 0 | 0 | 0 | 1 | 1 | 1 | 0 |
| **6. Study perspective** | 0.5 | 0.5 | 0.5 | 0 | 0.5 | 0.5 | 0.5 | 0.5 |
| **7. Comparator(s)** | 0.5 | 1 | 0.5 | 1 | 1 | 0.5 | 1 | 1 |
| **8. Time-horizon** | 0 | 0.5 | 0 | 0 | 1 | 0.5 | 1 | 1 |
| **9. Discount rate** | 0 | 0 | 1 | 0 | 1 | 0 | 0 | 1 |
| **10. Choice of health outcomes** | 0.5 | 0.5 | 0.5 | 0 | 1 | 1 | 0.5 | 0.5 |
| **11. Measurement of effectiveness** | 1 | 1 | 0 | 1 | 1 | 1 | 1 | 1 |
| **12. Measurement and validation of preference-based outcomes** | 1 | 0 | 0 | 0 | 1 | 0 | 0 | 1 |
| **13. Resources and costs** | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| **14. Currency, price data, and conversion** | 0 | 1 | 0.5 | 0 | 1 | 1 | 1 | 1 |
| **15. Choice of model** | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 |
| **16. Assumptions** | 0.5 | 0.5 | 0.5 | 0.5 | 1 | 1 | 1 | 0.5 |
| **17. Analytic methods** | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| **18. Study parameters** | 1 | 0.5 | 0.5 | 1 | 1 | 1 | 1 | 1 |
| **19. Incremental costs and outcomes** | 1 | 1 | 1 | 0.5 | 1 | 1 | 1 | 1 |
| **20. Uncertainty** | 0.5 | 1 | 1 | 0.5 | 1 | 1 | 1 | 1 |
| **21. Heterogeneity** | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 |
| **22. Discussion** | 1 | 1 | 0.5 | 1 | 1 | 1 | 1 | 1 |
| **23. Source of funding** | 1 | 1 | 1 | 0 | 1 | 1 | 1 | 0 |
| **24. Conflicts of interest** | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| **Total score** | **13.5** | **16** | **12.5** | **10.5** | **23** | **18.5** | **19** | **18.5** |

# Appendix L

# UNIVERSITY OF LEEDS – ASSIGNMENT OF STUDENT INTELLECTUAL PROPERTY RIGHTS

**PARTIES:**
Paola Cocco of 22 Brudenell Road, Leeds, LS6 1BD ("**Student**"); and
**THE UNIVERSITY OF LEEDS**, registered no. RC000658, of Leeds, LS2 9JT ("**University**")

It is agreed as follows:

1. With effect from 30th July 2020 and in consideration of the sum of 10 pence paid by the University (receipt of which the Student expressly acknowledges) and the opportunity to undertake further work in connection with the Activities the Student hereby assigns to the University absolutely all their rights, title and interest in the intellectual property rights defined in clause 2 below. To the extent that the Student is able, the Student furthermore hereby assigns all future IPR the Student may create in connection with the Activities.
2. For the purposes of this Agreement "**intellectual property rights**" or "**IPR**" means: patents, rights to inventions, copyright and related rights, trade marks, business names and domain names, rights in get-up, goodwill and the right to sue for passing off, rights in designs, database rights, rights to use, and protect the confidentiality of, confidential information (including know-how) and all other intellectual property rights, in each case whether registered or unregistered and including all applications and rights to apply for and be granted, renewals or extensions of, and rights to claim priority from, such rights and all similar or equivalent rights or forms of protection which subsist or will subsist now or in the future in any part of the world arising from any work or other contributions made by the Student in connection with the activity as defined under clause 3 below. For the avoidance of doubt, the Student does not waive any "moral rights" e.g the right to be acknowledged as author, (see https://www.gov.uk/guidance/the-rights-granted-by-copyright) in copyright works which may accrue in the Student's favour.
3. Activities means: Any original work being produced following access to COMBACTE EU CDI data for the project 'Developing a framework based on early economic evaluation to inform target product profiles for new diagnostic tests'. Access to COMBACTE EU CDI data will support the structure and parameterisation of a decision model on diagnostic tests for CDI. Specifically, COMBACTE EU CDI data will inform key input parameters in the decision model (e.g. sampling and testing rate, incidence and prevalence of CDI).
4. The University agrees to reward the Student if appropriate in accordance with the University's Intellectual Property Policy http://www.leeds.ac.uk/secretariat/documents/ipr_policy.pdf in the case of any commercial exploitation of the IPR assigned under this Agreement.
5. The Student warrants that the IPR assigned (or to be assigned) is free from any third party rights.
6. The Student agrees to do all reasonably required to ensure that the above IPR is fully assigned to the University. To the extent that the Student cannot legally assign future IPR by this Agreement, the Student undertakes to the University to enter into further assignments relating to future IPR in favour of the University.
7. This Agreement constitutes the entire understanding between the parties relating to the transfer of present and future IPR. Any variations to this Agreement must be made in writing.
8. This Agreement remains subject to English law and the non-exclusive jurisdiction of the English Courts.

**SIGNED** by the Student Paola Cocco Date: 30th July 2020

**SIGNED** for and on behalf of the University of Leeds: ..... Date: 5th Aug 2020

Name: **Peter Muller**

# Appendix M

# Reporting checklists – early economic model for new hypothetical rapid diagnostic test for CDI

For each item of the STRESS-DES and CHEERS checklists, this appendix provides a reference to the thesis chapter(s) and section(s) where to find additional information on the development of the model outlined in Chapter 5 and Chapter 6.

### Table M-1 CHEERS checklist (286)

| Topic | No. | Item | Location where item is reported |
|---|---|---|---|
| **Title** | | | |
| | 1 | Identify the study as an economic evaluation and specify the interventions being compared. | Chapter 5, Title |
| **Abstract** | | | |
| | 2 | Provide a structured summary that highlights context, key methods, results, and alternative analyses. | Not applicable |
| **Introduction** | | | |
| **Background and objectives** | 3 | Give the context for the study, the study question, and its practical relevance for decision making in policy or practice. | Chapter 5, section 5.3.1. |
| **Methods** | | | |
| **Health economic analysis plan** | 4 | Indicate whether a health economic analysis plan was developed and where available. | Not applicable |
| **Study population** | 5 | Describe characteristics of the study population (such as age range, demographics, socioeconomic, or clinical characteristics). | Chapter 5, section 5.3.3.2. |
| **Setting and location** | 6 | Provide relevant contextual information that may influence findings. | Chapter 5, section 5.3.3, section 5.3.3.3, section 5.3.3.5-5.3.3.7. |
| **Comparators** | 7 | Describe the interventions or strategies being compared and why chosen. | Chapter 5, section 5.3.3.4.1-5.3.3.4.2. |
| **Perspective** | 8 | State the perspective(s) adopted by the study and why chosen. | Chapter 5, section 5.3.1, section 5.3.6. |
| **Time horizon** | 9 | State the time horizon for the study and why appropriate. | Chapter 5, section 5.3.7.1 |
| **Discount rate** | 10 | Report the discount rate(s) and reason chosen. | Chapter 5, section 5.3.6. |
| **Selection of outcomes** | 11 | Describe what outcomes were used as the measure(s) of benefit(s) and harm(s). | Chapter 5, section 5.3.6. |

| Topic | No. | Item | Location where item is reported |
|---|---|---|---|
| **Measurement of outcomes** | 12 | Describe how outcomes used to capture benefit(s) and harm(s) were measured. | Chapter 5, section 5.3.6. |
| **Valuation of outcomes** | 13 | Describe the population and methods used to measure and value outcomes. | Chapter 5, section 5.3.5.9. |
| **Measurement and valuation of resources and costs** | 14 | Describe how costs were valued. | Chapter 5, section 5.3.5.10. |
| **Currency, price date, and conversion** | 15 | Report the dates of the estimated resource quantities and unit costs, plus the currency and year of conversion. | Chapter 5, section 5.3.5.10. |
| **Rationale and description of model** | 16 | If modelling is used, describe in detail and why used. Report if the model is publicly available and where it can be accessed. | Chapter 5, section 5.3.2, section 5.3.3. |
| **Analytics and assumptions** | 17 | Describe any methods for analysing or statistically transforming data, any extrapolation methods, and approaches for validating any model used. | Chapter 5, section 5.3.3.1, section 5.3.5.1, section 5.3.8. |
| **Characterising heterogeneity** | 18 | Describe any methods used for estimating how the results of the study vary for subgroups. | Not applicable |
| **Characterising distributional effects** | 19 | Describe how impacts are distributed across different individuals or adjustments made to reflect priority populations. | Not applicable |
| **Characterising uncertainty** | 20 | Describe methods to characterise any sources of uncertainty in the analysis. | Chapter 5, section 5.3.7.2. Chapter 6, section 6.2.2-6.2.3, |
| **Approach to engagement with patients and others affected by the study** | 21 | Describe any approaches to engage patients or service recipients, the general public, communities, or stakeholders (such as clinicians or payers) in the design of the study. | Chapter 5, section 5.3.3. |
| **Results** | | | |
| **Study parameters** | 22 | Report all analytic inputs (such as values, ranges, references) including uncertainty or distributional assumptions. | Chapter 5, Table 5-4, section 5.3.5. |
| **Summary of main results** | 23 | Report the mean values for the main categories of costs and outcomes of interest and summarise them in the most appropriate overall measure. | Chapter 6, Table 6-3, Table 6-4 |
| **Effect of uncertainty** | 24 | Describe how uncertainty about analytic judgments, inputs, or projections affect findings. Report the effect of choice of discount rate and time horizon, if applicable. | Chapter 6, section 6.3.2-6.3.3 |
| **Effect of engagement with patients and others affected by the study** | 25 | Report on any difference patient/service recipient, general public, community, or stakeholder involvement made to the approach or findings of the study | Not applicable |
| **Discussion** | | | |
| **Study findings, limitations, generalisability, and current knowledge** | 26 | Report key findings, limitations, ethical or equity considerations not captured, and how these could affect patients, policy, or practice. | Chapter 6, section 6.4.1, section 6.4.3 |

| Topic | No. | Item | Location where item is reported |
|---|---|---|---|
| **Other relevant information** | | | |
| **Source of funding** | 27 | Describe how the study was funded and any role of the funder in the identification, design, conduct, and reporting of the analysis | Page iv |
| **Conflicts of interest** | 28 | Report authors conflicts of interest according to journal or International Committee of Medical Journal Editors requirements. | Not applicable |

**Table M-2 STRESS-DES guidelines checklist (285)**

| Section/Subsection | Item | Reference to the thesis |
|---|---|---|
| **1. Objectives** | | |
| Purpose of the model | 1.1 | See Chapter 5 section, 5.3.1 |
| Model Outputs | 1.2 | See Chapter 5, section 5.3.3.7 |
| Experimentation Aims | 1.3 | Scenario-based analysis – see Chapter 6, section 6.2.2 |
| **2. Logic** | | |
| Base model overview diagram | 2.1 | See Chapter 5, section 5.3.3 |
| Base model logic | 2.2 | See Chapter 5, section 5.3.3 |
| Scenario logic | 2.3 | See Chapter 6, section 6.2.2 |
| Algorithms | 2.4 | See Technical documentation - available upon request |
| Components | 2.5 | See Chapter 5, section 5.3.3.1- 5.3.3.2 |
| | | See Chapter 5, section 5.3.3 |
| | | See Chapter 5, section 5.3.3.2- 5.3.3.3- 5.3.3.4- 5.3.3.7 |
| | | See Chapter 5, section 5.3.3.1 |
| **3. Data** | | |
| Data sources | 3.1 | See Chapter 5, section 5.3.5. |
| Pre-processing | 3.2 | See Chapter 5, section 5.3.5. Appendix N |
| Input parameters | 3.3 | See Chapter 5, section 5.3.5, Appendix N |
| Assumptions | 3.4 | See Chapter 5, section 5.3.3.1. |
| **4. Experimentation** | | |
| Initialisation | 4.1 | See Chapter 5, section 5.3.7.1, Appendix N |
| Run length | 4.2 | See Chapter 5, section 5.3.7 |
| Estimation approach | 4.3 | See Chapter 5, section 5.3.7.1, Appendix N |
| **5. Implementation** | | |
| Software or programming language | 5.1 | See Chapter 5, section 5.3.7 |
| Random sampling | 5.2 | See Chapter 5, section 5.3.7.1 |
| Model execution | 5.3 | See Chapter 5, section 5.3.7 |
| System Specification | 5.4 | See Chapter 6, section 6.2. |
| **6. Code Access** | | |
| Computer Model Sharing Statement | 6.1 | See Chapter 5, section 5.3.3. |

# Appendix N
# Model parametrisation

This appendix provides additional information on the parameterisation of the model described in Chapter 5 and Chapter 6. In particular, further information is provided on: (i) how the inter-arrival time is estimated within the model (see section N.1); (ii) how the simulated availability of single room is calculated (section N.2); and (iii) results from the parametric survival analysis on time-to-event variables in the model (i.e. duration of symptoms for CDI positive and negative patients, and hospital LOS) (section N.3).

## N.1 Inter-arrival time calculations

**Table N-1 Monthly average stool samples being tested in the UK, sorted by hospital type (source: COMBACTE-CDI survey)**

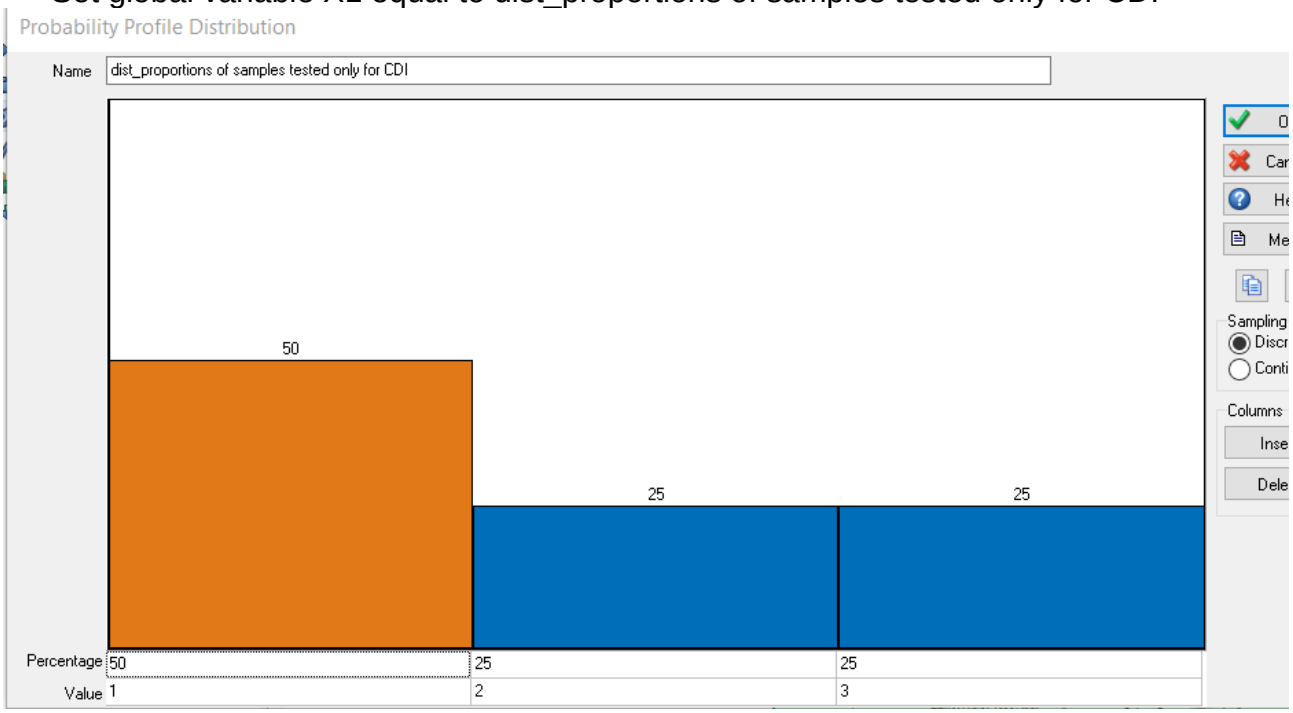
| Monthly average stool samples being tested | | |
|---|---|---|
| | ***District Hospitals*** | ***Teaching Hospitals*** |
| Mean | 984 | 1923 |
| Standard Error | 236 | 688.7448 |
| Median | 850 | 1430.5 |
| Standard Deviation | 709.324111 | 1948.065 |
| Minimum | 162 | 70 |
| Maximum | 2425 | 6239 |
| Count | 9 | 8 |
| Confidence Level (95%) | 545.234778 | 1628.623 |
| Upper CI (95%) | 1529 | 3551 |
| Lower CI (95%) | 439 | 294 |

**Table N-2 Percentage of stool samples tested only for CDI in UK hospitals (source: COMBACTE-CDI survey)**

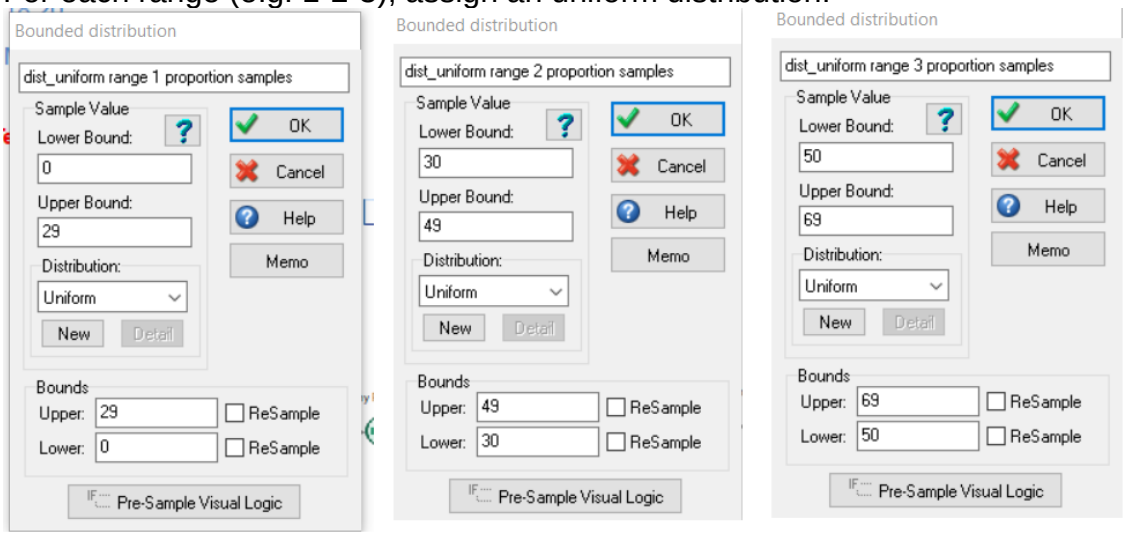
| | n | Frequency |
|---|---|---|
| Less than 30% of samples | 2 | 0.25 |
| 49-30% of samples | 4 | 0.5 |
| 69-50% of samples | 2 | 0.25 |

The following steps were taken to calculate the inter-arrival time, including:

1. Calculate the proportion of samples being tested for CDI only
   - Set global variable X1 equal to dist_proportions of samples tested only for CDI




     - Value 1: 30-49% of samples being tested for CDI only
     - Value 2: less than 30% of samples being tested for CDI only
     - Value 3: 50-69% of samples being tested for CDI only
   - Assume an equal probability within each range (e.g. if you consider value 2, 0-10%, 10-20%, 20-30% ranges have a probability of 0.33). This is because COMBACTE-CDI survey data does not report the frequency within each range. For each range (e.g. 1-2-3), assign an uniform distribution.



```
SET X1 = dist_proportions of samples tested only for CDI
IF X1 = 1
    SET X2 = dist_equal probability
    IF X2 = 1
        SET p_sample tested CDI only = 0.3
    ELSE IF X2 = 2
        SET p_sample tested CDI only = 0.4
    ELSE IF X2 = 3
        SET p_sample tested CDI only = 0.49
ELSE IF X1 = 2
    SET X2 = dist_equal probability
    IF X2 = 1
        SET p_sample tested CDI only = 0.1
    ELSE IF X2 = 2
        SET p_sample tested CDI only = 0.2
    ELSE IF X2 = 3
        SET p_sample tested CDI only = 0.29
ELSE IF X1 = 3
    SET X2 = dist_equal probability
    IF X2 = 1
        SET p_sample tested CDI only = 0.5
    ELSE IF X2 = 2
        SET p_sample tested CDI only = 0.6
    ELSE IF X2 = 3
        SET p_sample tested CDI only = 0.69
```

```
Monthly average samples VL
    1. Calculate the proportion of samples being tested for CDI only
    SET X1 = dist_ range proportions of samples tested for CDI only
    IF X1 = 1
        SET p_sample tested CDI only = [dist_uniform range 1 proportion samples]/100
    ELSE IF X1 = 2
        SET p_sample tested CDI only = [dist_uniform range 2 proportion samples]/100
    ELSE IF X1 = 3
        SET p_sample tested CDI only = [dist_uniform range 3 proportion samples]/100
```

2. Calculate monthly average of samples tested only for CDI
    - Depending on the hospital size, multiply the median stools being tested with the proportion of samples being tested only for CDI to monthly average samples
    - Calculate weekly average= monthly average CDI only/4
    - Calculate daily average= weekly average/5 working days
    - Calculate number of samples per minute= daily average/600 min[36]
    - Lambda= 1/number of samples per minute
    - Plug lambda in an exponential distribution

[36] The simulation runs from 9am to 7pm (10 hours) on Monday-Friday. Each day has 60*10= 600 minutes

```
SET Monthly average stools = dist_average stool samples_Teaching Hospital
SET Monthly average stools tested CDI only = p_sample tested CDI only*Monthly average stools
SET Weekly average stools = Monthly average stools tested CDI only/4
SET Daily average stools = Weekly average stools/5
SET X3 = Daily average stools/600
SET lambda = 1/X3
```

- Lambda = 1/ number of samples per minute

3. Set Start Point inter-arrival time equal to an exponential distribution with lambda as average

## N.2 Single room availability

Assumption - equilibrium in the other hospital wards. Take a 1/3 of the average of single rooms available. The following estimates were sense-checked with a clinical expert.

- District Hospitals: 196.8/3= 66 single rooms
- Teaching Hospitals: 278.7/3= 93 single rooms

**Table N-3 Single rooms availability in UK hospitals, sorted by hospital type (source: COMBACTE-CDI survey)**

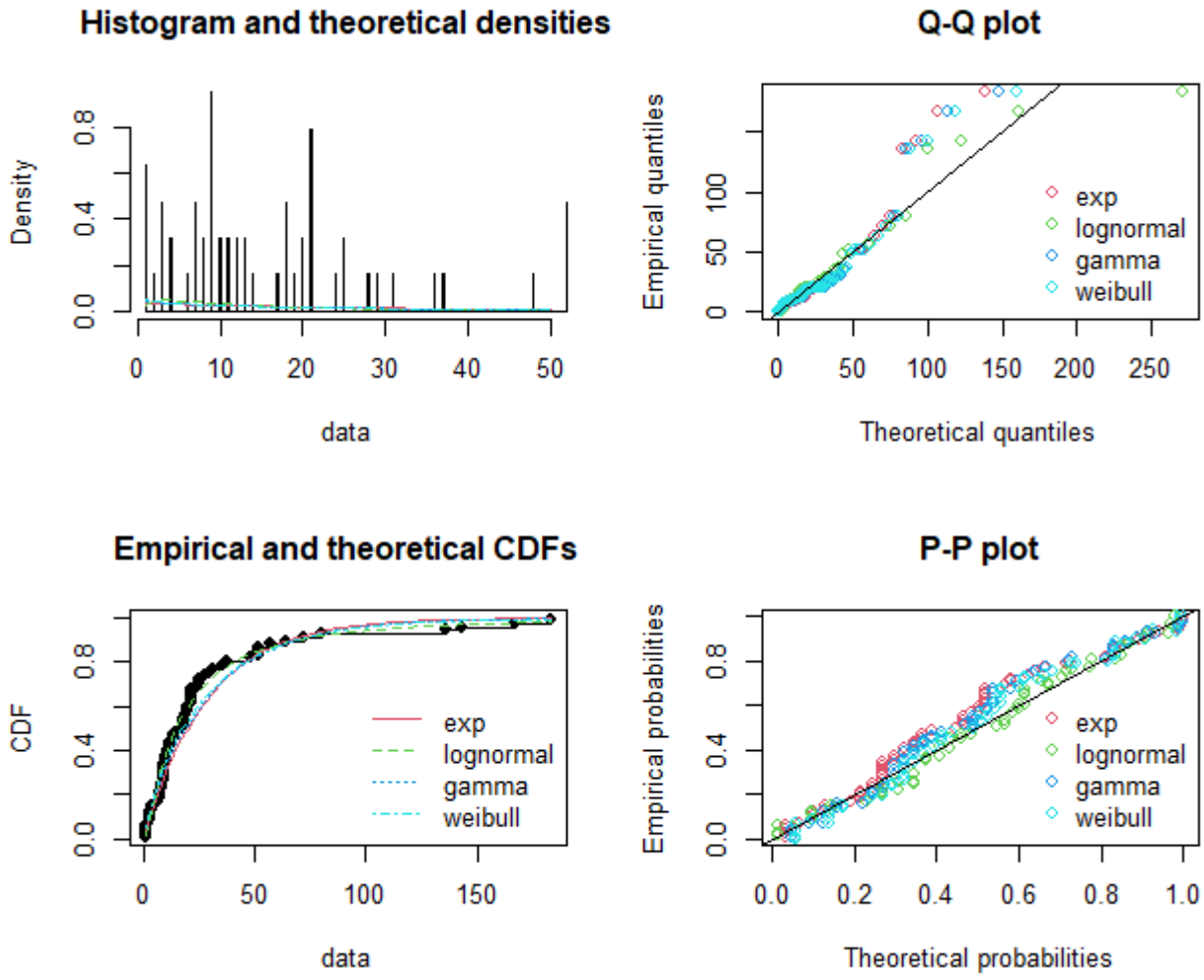
| | District Hospitals | Teaching Hospitals |
|---|---|---|
| Mean | 196.8 | 278.7 |
| Standard Deviation | 108.6 | 93.8 |
| Upper bound 95% CI | 280.3 | 365.5 |
| Lower bound 95% CI | 113.4 | 192 |

## N.3 Parametric survival analysis

### N.3.1 Visual inspection of goodness of fit of different parametric distribution against the observed IPD

**LOS CDI negative**

**Figure N-1 Histogram and theoretical densities, Q-Q plot, empirical and theoretical Cumulative Distribution Functions (CDFs), and P-P plot for different parametric distributions simulating LOS CDI negative IPD**



**LOS CDI positive**

***Slow time-to-result***

**Figure N-2 Histogram of fitted Weibull distribution to LOS CDI positive IPD (slow time-to-result)**

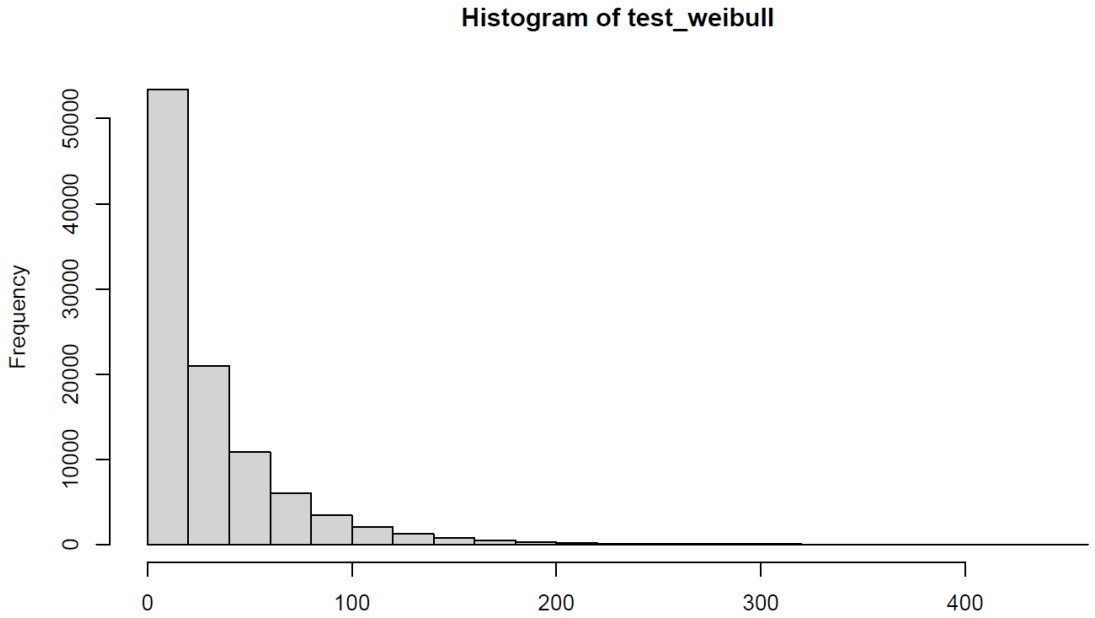



**Figure N-3 Histogram of fitted exponential distribution to LOS CDI positive IPD (slow time-to-result)**

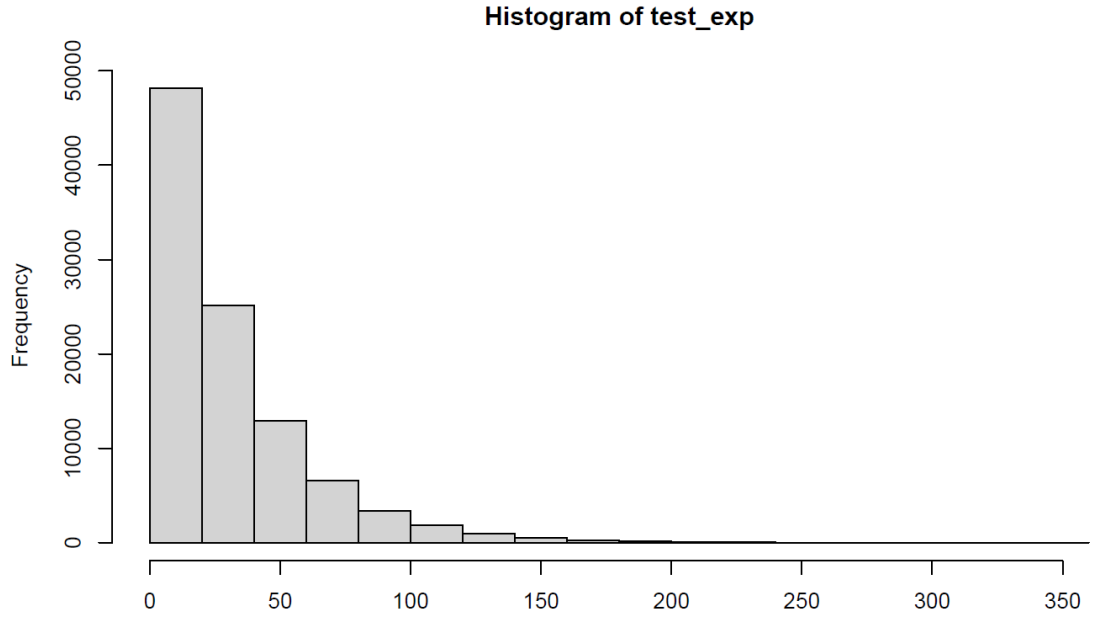
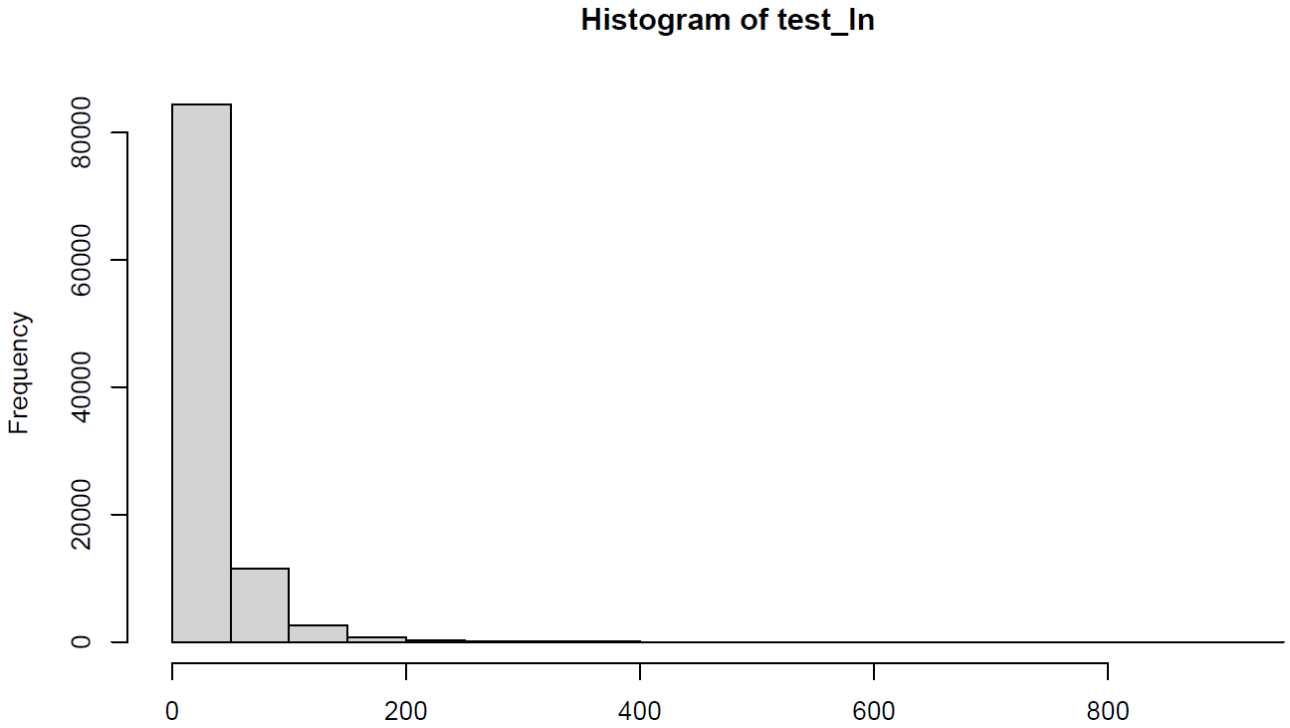


**Figure N-4 Histogram of fitted lognormal distribution to LOS CDI positive IPD (slow time-to-result)**




***Average time-to-diagnosis***

**Figure N-5 Histogram of fitted Weibull distribution to LOS CDI positive IPD (average time-to-result)**

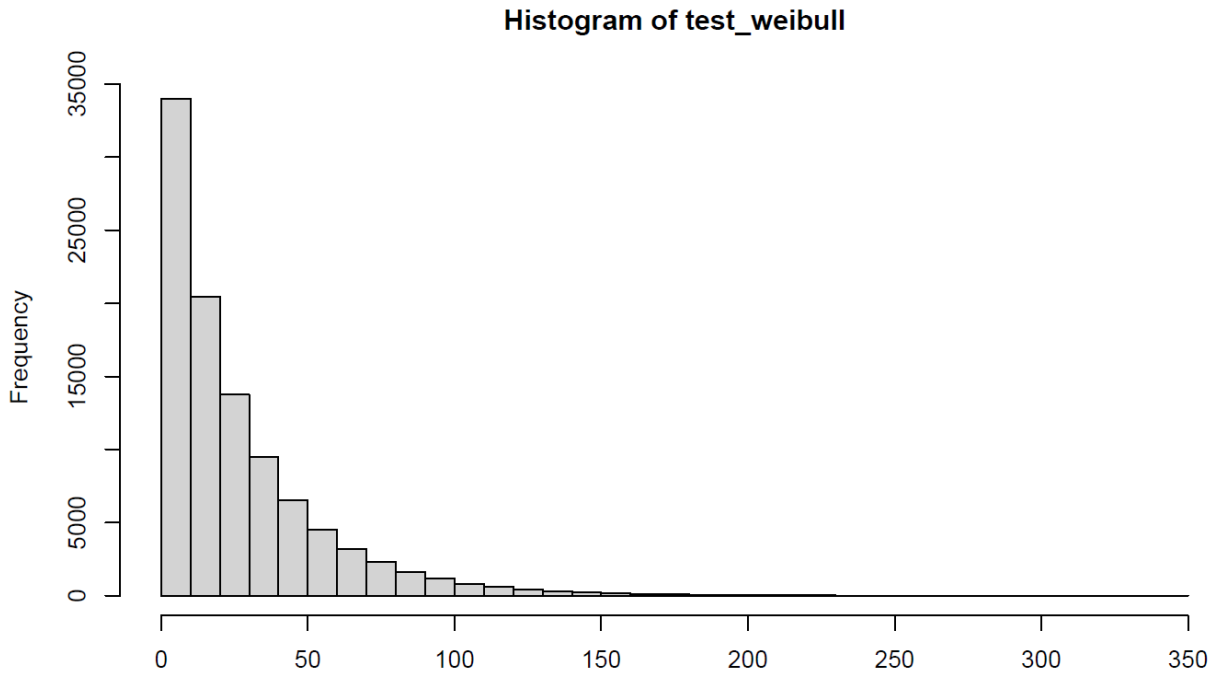


**Figure N-6 Histogram of fitted exponential distribution to LOS CDI positive IPD (average time-to-result)**

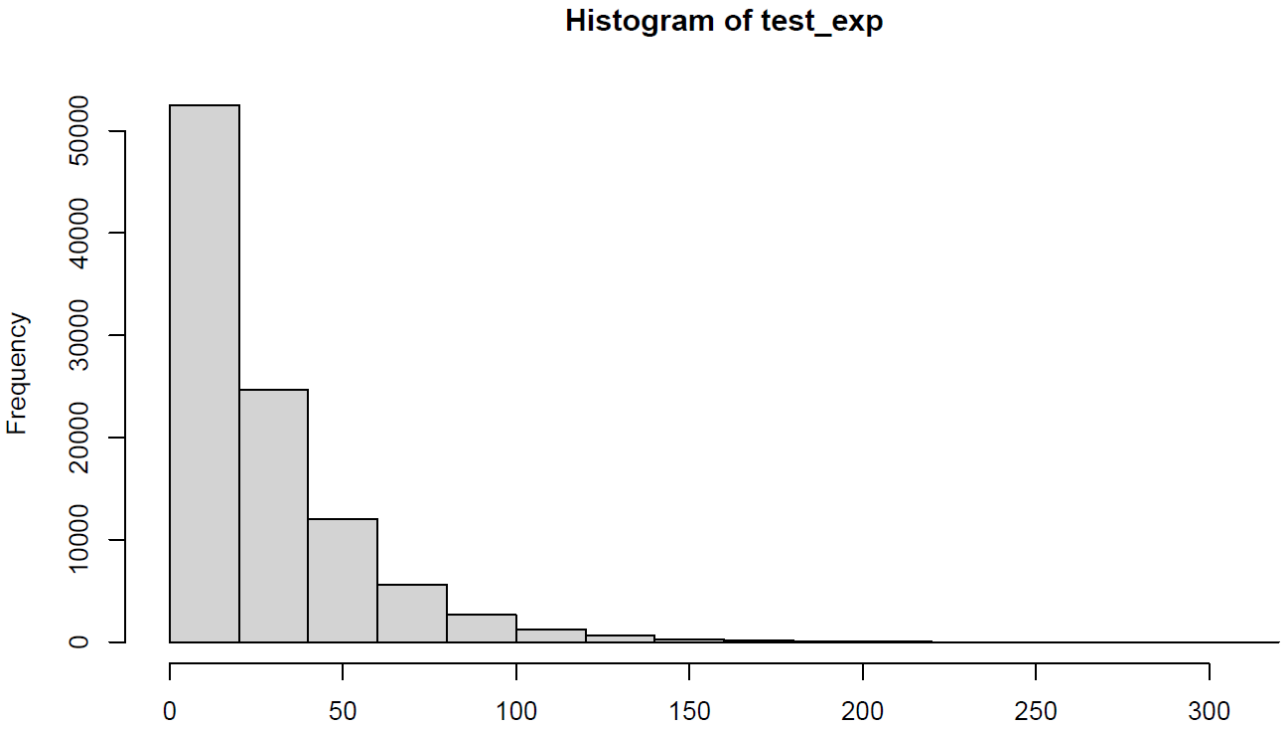
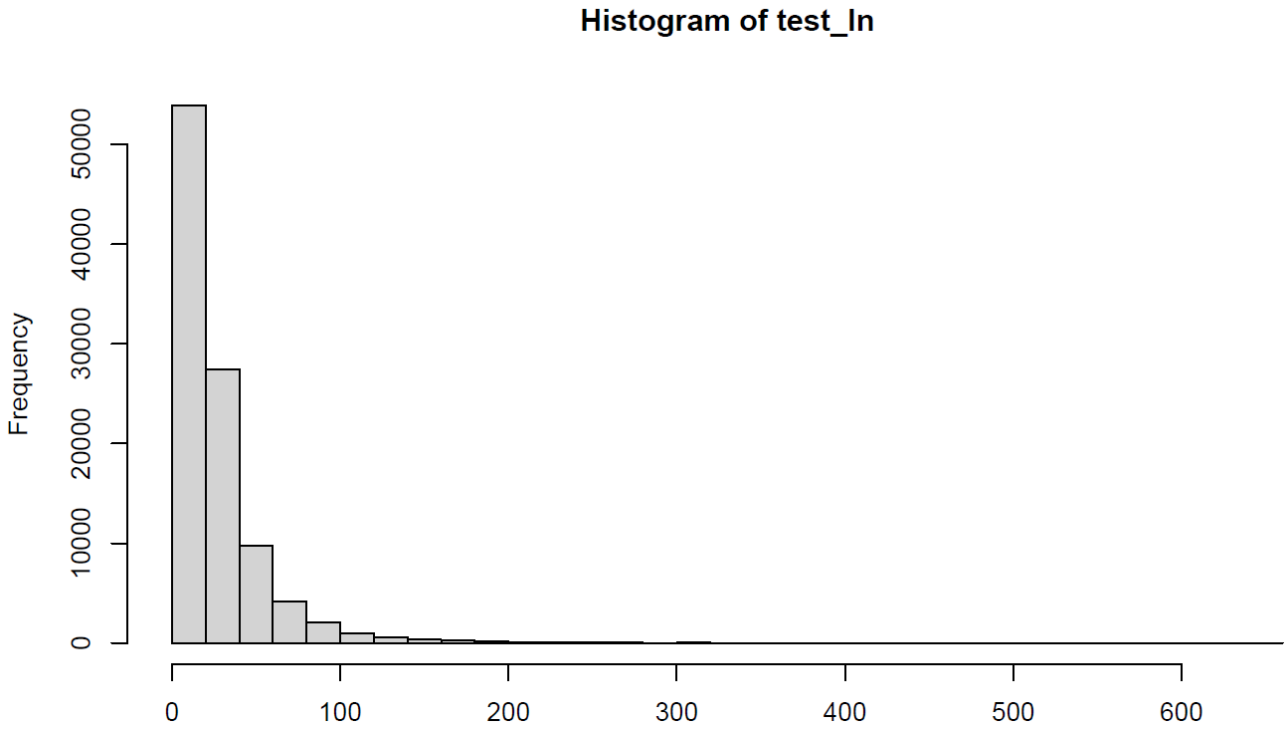



**Figure N-7 Histogram of fitted lognormal distribution to LOS CDI positive IPD (average time-to-result)**



***Rapid time-to-diagnosis***

**Figure N-8 Histogram of fitted Weibull distribution to LOS CDI positive IPD (rapid time-to-result)**

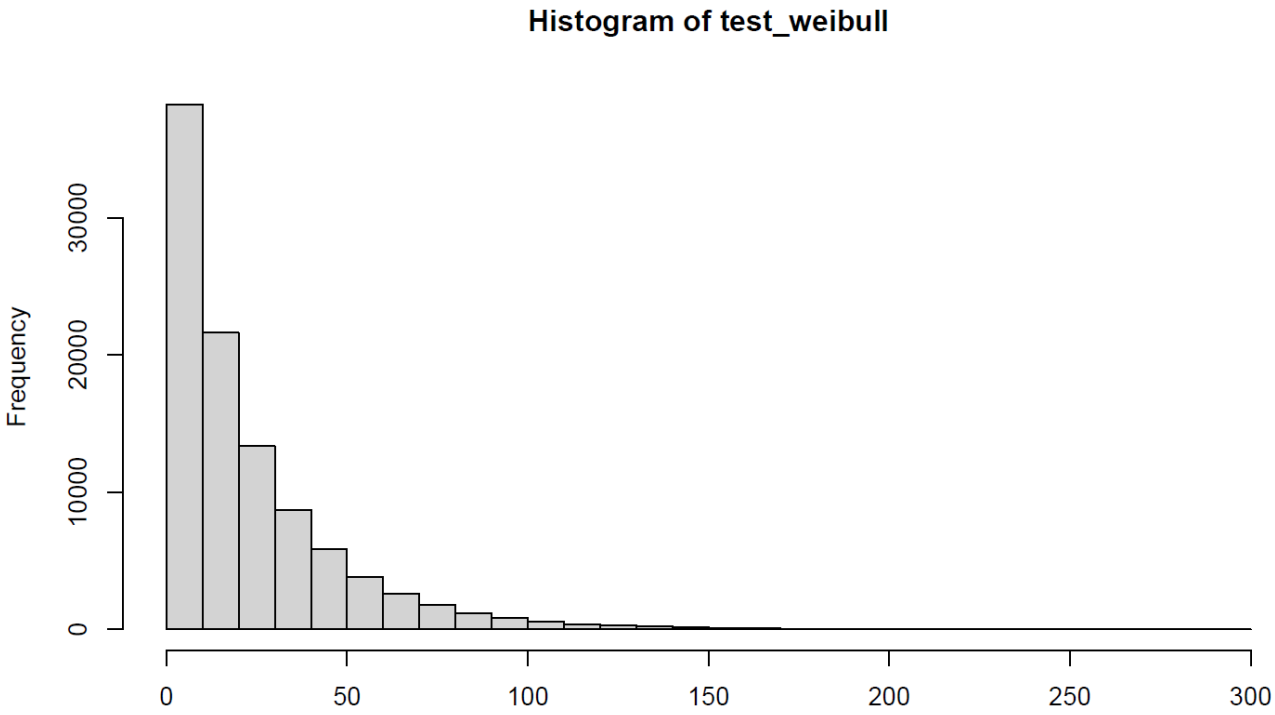



**Figure N-9 Histogram of fitted exponential distribution to LOS CDI positive IPD (rapid time-to-result)**

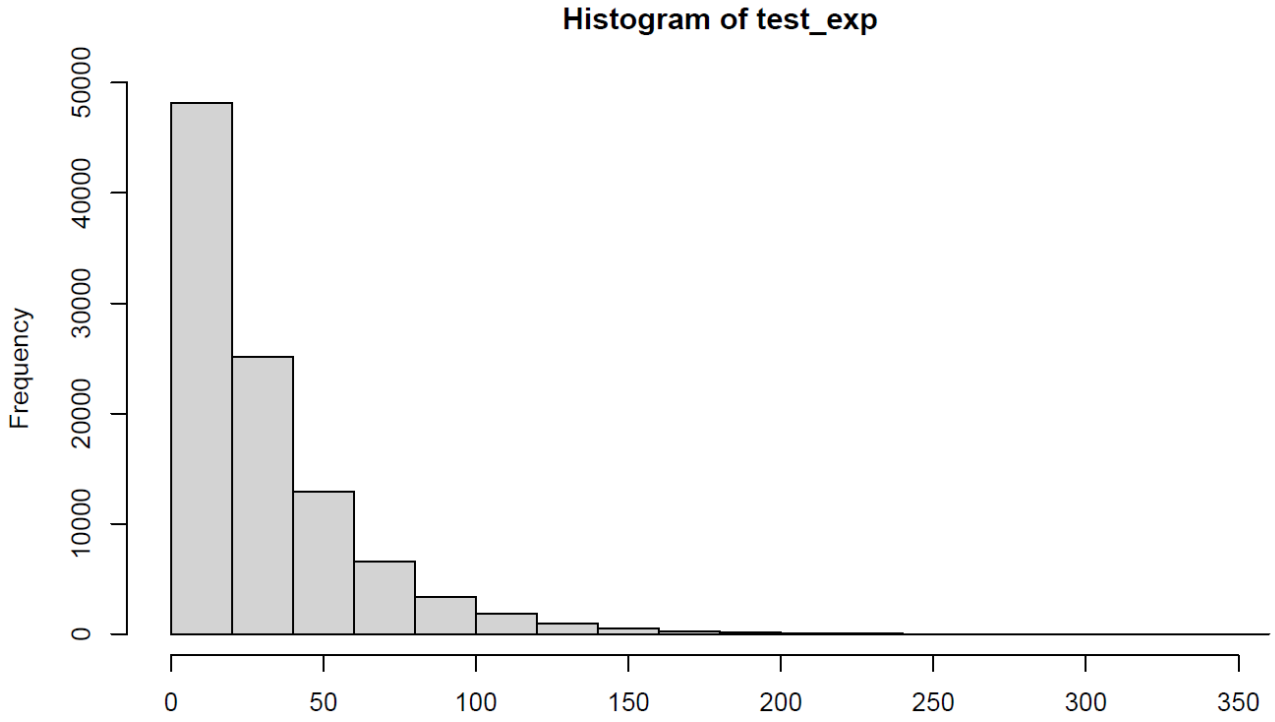
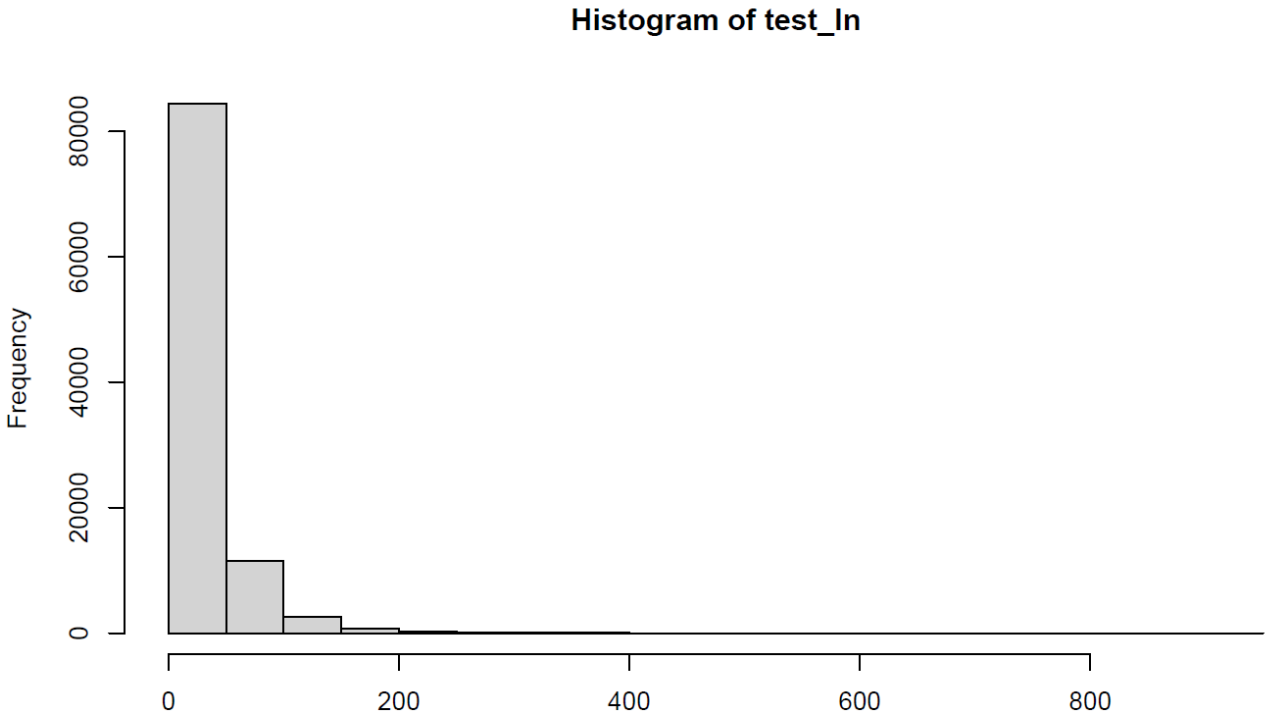
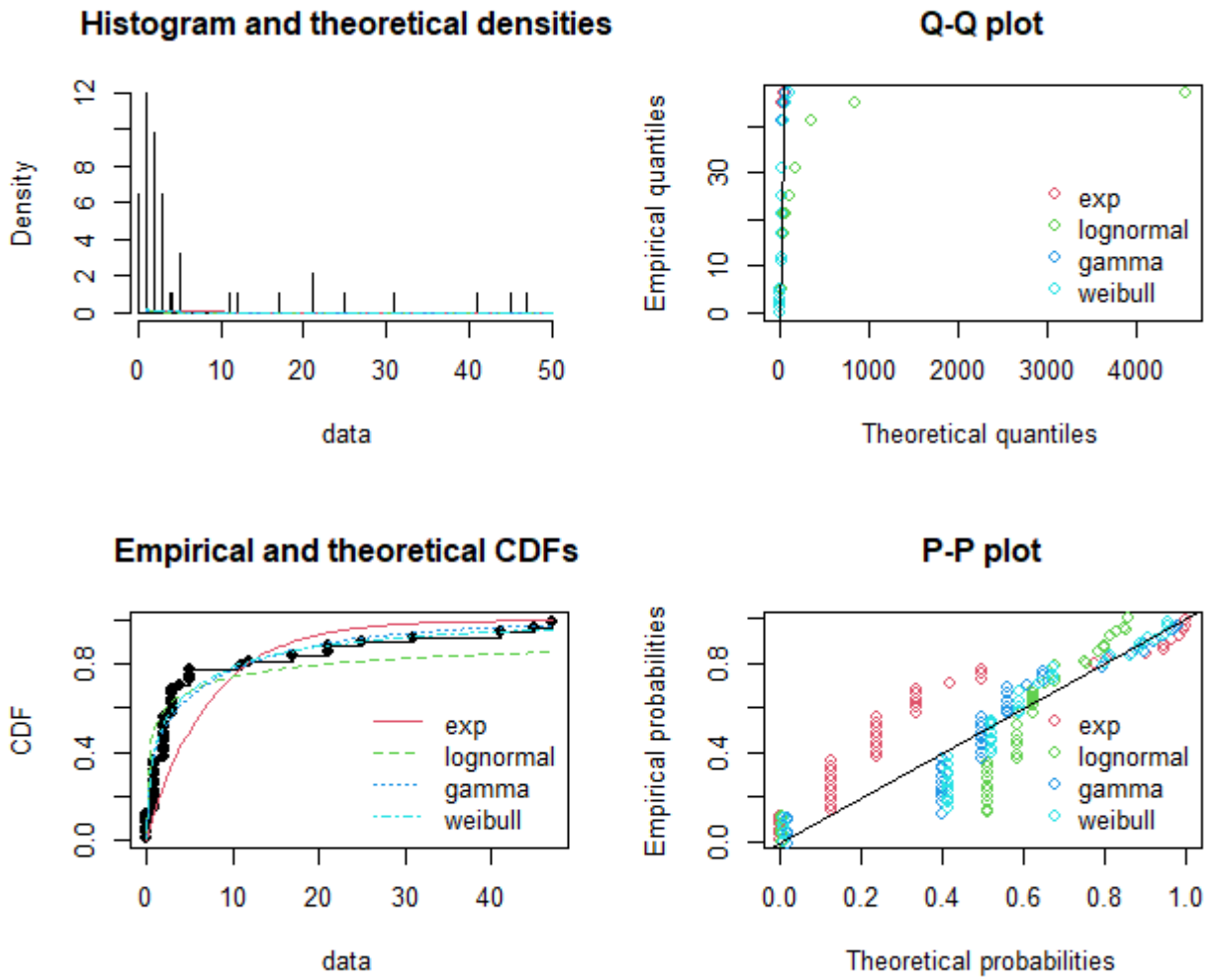


**Figure N-10 Histogram of fitted lognormal distribution to LOS CDI positive IPD (rapid time-to-result)**



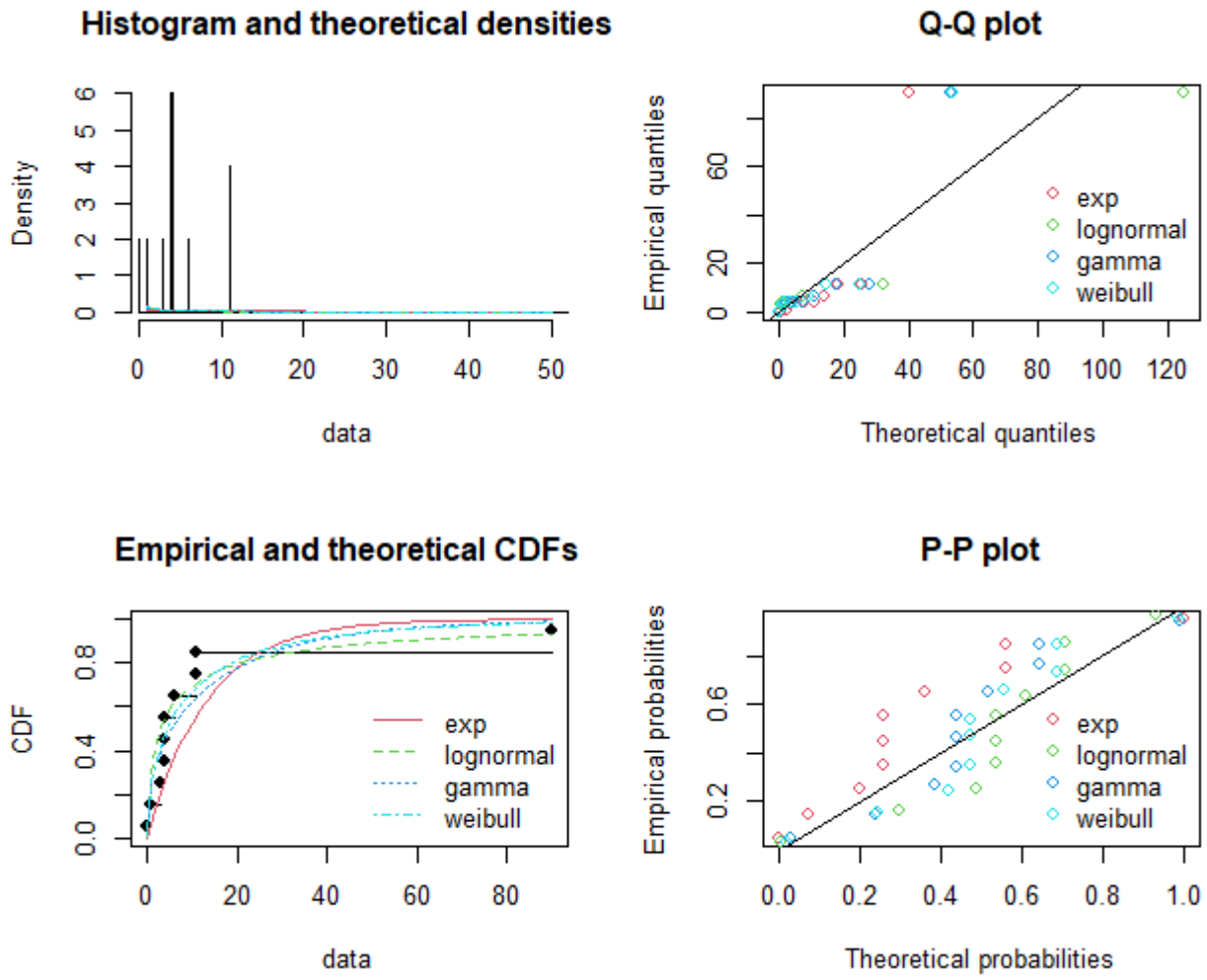
**Diarrhoea duration CDI negative**

**Figure N-11 Histogram and theoretical densities, Q-Q plot, empirical and theoretical CDFs, and P-P plot for different parametric distributions simulating diarrhoea duration CDI negative IPD**



**Diarrhoea duration CDI positive**

**Figure N-12 Histogram and theoretical densities, Q-Q plot, empirical and theoretical CDFs, and P-P plot for different parametric distributions simulating diarrhoea duration CDI positive IPD**



## Example R code for parametric sampling method

```
#Author: Alison Smith
#set wd
setwd('')   #Change to link to your working directory where the data is stored
# Clear the workspace
rm(list = ls());
# Load packages
install.packages("MASS")
install.packages("survival")
install.packages('fitdistrplus')
library(MASS)
library(survival)
library(fitdistrplus)
library(MASS)
# Load dataset
df_data <- read.csv(file = "")   #Change to according to the file name
#============== Survival analysis ====
#Make sure LOS data is in numeric format
#e.g. if data column is called "LOS":
data_temp <- as.numeric(df_data$ )   #change variable name here as required
length(data_temp[is.na(data_temp)])
summary(data_temp)
#zero values cause an issue so add on a small amount to zero values if there are any
#data_temp <- ifelse(df_data==0, 0.1, df_data)
#summary(data_temp)
length(data_temp[data_temp<0])      #check if you have any negative values (need to deal with them if so)
length(data_temp[is.na(data_temp)])    #check if you have any NA values   (need to deal with them if so)
data_temp <- data_temp[!is.na(data_temp)]#here I guess I am removing NA values
length(data_temp[data_temp<0]) ##now there are no negative values
data_temp <- ifelse(data_temp==0,0.01,data_temp)
length(data_temp[data_temp<0])
#Plots to explore possible fits
#indicates which distributions are worth considering
plotdist(data_temp, histo=TRUE, demp=TRUE)
descdist(data_temp, boot=1000)
#Fit distributions
fit_exp   <- fitdist(data=data_temp, distr="exp", method="mle")
fit_lnorm <- fitdist(data=data_temp, distr="lnorm", method="mle")
fit_gamma <- fitdist(data=data_temp, distr="gamma", method="mle")
fit_weib  <- fitdist(data=data_temp, distr="weibull", method="mle")
```

**#Compare goodness of fits**

```
fit_exp$aic; fit_lnorm$aic; fit_gamma$aic; fit_weib$aic

fit_exp$bic; fit_lnorm$bic; fit_gamma$bic; fit_weib$bic
```

**# #Plots**

```
par(mfrow = c(2, 2))  #sets plots to show 2x2 layout

plot.legend <- c("exp", "lognormal", "gamma", "weibull")

denscomp(list(fit_exp, fit_lnorm, fit_gamma, fit_weib), xlim = c(0,10),addlegend = TRUE, legendtext = plot.legend) #what is this plot showing?

legend(legend = c("exp", "lognormal", "gamma", "weibull"))

qqcomp(list(fit_exp, fit_lnorm, fit_gamma, fit_weib), legendtext = plot.legend)

cdfcomp(list(fit_exp, fit_lnorm, fit_gamma, fit_weib), legendtext = plot.legend)

ppcomp(list(fit_exp, fit_lnorm, fit_gamma, fit_weib), legendtext = plot.legend)

par(mfrow = c(1, 1)) #resets plot layout to a single plot
```

**#Check the fits against observed data**

```
summary(rlnorm(10000,fit_lnorm$estimate[1], fit_lnorm$estimate[2]))

summary(rweibull(10000,fit_weib$estimate[1], fit_weib$estimate[2]))

summary(rgamma(10000,fit_gamma$estimate[1], fit_gamma$estimate[2]))

summary(rexp(10000,fit_exp$estimate[1]))

summary(data_temp)
```

**#Get parameter values to use in model**

```
summary(fit_exp)

summary(fit_gamma)

summary(fit_lnorm)

summary(fit_weib)
```

**####FITTING DISTRIBUTIONS FOR PARAMETERS W/O INDIVIDUAL PATIENT DATA**

```
library(mixdist)

mean_slow_TAT <- 30.3 ## LOS CDI Pos patients with slow time-to-result - change this as appropriate

sd_slow_TAT <- 36.3

mean_rapid_TAT <- 23.2

sd_rapid_TAT <- 25.4

mean_average_TAT <- 26.9

sd_average_TAT <- 28.9

## 1- WEIBULL DISTRIBUTION

weibullpar(mean_rapid_TAT,sd_rapid_TAT)

##this gives shape 0.8389166 and scale 27.61969

test_weibull <-rweibull(100000, shape=0.9145084, scale = 22.23937)

hist(test_weibull,breaks=30)

summary(test_weibull)

## 2- LOGNORMAL DISTRIBUTION

meanlog <- log(mean_rapid_TAT^2 / sqrt(sd_rapid_TAT^2 + mean_rapid_TAT^2))    ##location parameter = meanlog
```

```
sdlog <- sqrt(log(1 + (sd_rapid_TAT^2 / mean_rapid_TAT^2)))## shape parameter = sdlog
print(meanlog)
print(sdlog)
test_ln <-rlnorm(100000,meanlog = meanlog, sdlog = sdlog)
hist(test_ln, breaks = 30)
summary(test_ln)
## 3- EXPONENTIAL DISTRIBUTION
lambda <- 1/mean_rapid_TAT   #rate parameter
print(lambda)
test_exp <- rexp(100000, rate = lambda)
summary(test_exp)
hist(test_exp)
length(test_weibull[which(test_weibull>100)])
length(test_ln[which(test_ln>100)])
length(test_exp[which(test_exp>100)])
length(test_weibull[which(test_weibull>200)])
length(test_ln[which(test_ln>200)])
length(test_exp[which(test_exp>200)])
```

### N.3.2 Analysis of the AIC and BIC metrics and the key summary statistics produced from different parametric distributions

**Table N-4 AIC, BIC criteria and descriptive statistics for the LOS and diarrhoea duration parametric distributions**

| | LOS CDI NEGATIVE | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | **AIC** | **BIC** | **Min** | **1st Quantile** | **Median** | **Mean** | **3rd Quantile** | **Max** | **Comments** |
| **Original dataset** | | | **1** | **8.5** | **17** | **28.6** | **28.5** | **183** | |
| Exponential | 550.5432 | 552.6864 | 0.0001 | 8.2424 | 19.885 | 28.1944 | 38.8163 | 360.192 | Second best choice |
| Lognormal | 545.1519 | 549.4382 | 0.2041 | 6.5302 | 14.5679 | 30.938 | 33.8305 | 1218.1461 | Although this distribution has the lowest AIC and BIC scores, descriptive statistics are not a closer fit and maximum value is too high (i.e. long tail) |
| Gamma | 551.8314 | 556.1176 | 0 | 6.935 | 18.513 | 28.323 | 39.553 | 382.521 | Second highest AIC and BIC scores |
| ***Weibull*** | 550.4396 | 554.7259 | 0.00038 | 6.55127 | 17.33442 | 28.49358 | 39.18863 | 302.83974 | Best choice given graphs and fitting with median and mean, AIC and BIC scores |
| | **Diarrhoea duration CDI NEGATIVE** | | | | | | | | |
| | **AIC** | **BIC** | **Min** | **1st Quantile** | **Median** | **Mean** | **3rd Quantile** | **Max** | **Comments** |
| **Original dataset** | | | **0.01** | **1** | **2** | **7.043** | **5** | **47** | |
| Exponential | 285.397 | 287.2682 | 0.00164 | 2.05785 | 4.90652 | 7.02925 | 9.70991 | 66.4793 | Highest AIC and BIC scores |
| Lognormal | 259.7221 | 263.4645 | 0 | 0.291 | 1.45 | 24.654 | 7.053 | 14068.535 | |
| Gamma | 251.3209 | 255.0633 | 0 | 0.4531 | 2.6035 | 6.8456 | 8.6629 | 129.0019 | Second highest AIC and BIC scores |
| ***Weibull*** | 250.0606 | 253.803 | 0 | 0.4446 | 2.1062 | 7.238 | 7.7828 | 264.8665 | Lowest AIC and BIC, closer to median, mean, 3rd quantile. Maximum value is high |

| | Diarrhoea duration CDI POSITIVE | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | **AIC** | **BIC** | **Min** | **1st Quantile** | **Median** | **Mean** | **3rd Quantile** | **Max** | **Comments** |
| **Original dataset** | | | **0.01** | **2** | **4** | **12.27** | **8.5** | **90** | |
| Exponential | 79.16398 | 79.56187 | 0.00175 | 3.34671 | 8.31677 | 12.01561 | 16.66562 | 122.07254 | Highest AIC and BIC scores |
| Lognormal | 75.48467 | 76.28046 | 0.001 | 0.652 | 2.777 | 23.922 | 11.734 | 4861.523 | Second highest AIC and BIC scores, maximum value is too high |
| Gamma | 74.43851 | 75.2343 | 0 | 0.9589 | 5.138 | 12.4303 | 16.3026 | 209.5033 | |
| ***Weibull*** | 73.3838 | 74.1796 | 0 | 0.9474 | 3.9532 | 11.3472 | 13.109 | 307.5213 | Weibull has the lowest AIC and BIC scores, although the max value is too high. |
| | **LOS CDI POSITIVE** | | | | | | | | |
| | **AIC** | **BIC** | **Min** | **1st Quantile** | **Median** | **Mean** | **3rd Quantile** | **Max** | **Comments** |
| **Original dataset** | | | **1** | **4** | **7** | **22** | **40** | **78** | |
| Exponential | 141.0954 | 141.9287 | 0.0032 | 6.3885 | 15.1652 | 22.1214 | 30.9259 | 240.3646 | |
| Lognormal | 140.472 | 142.1384 | 0.0258 | 3.7077 | 9.6051 | 24.5721 | 24.5303 | 1347.278 | Although this distribution has the lowest AIC and BIC scores, descriptive statistics are not a closer fit and maximum value is too high (i.e. long tail) |
| Gamma | 141.8438 | 143.5103 | 0.00021 | 4.41184 | 13.32876 | 22.06313 | 29.89258 | 253.45255 | |
| ***Weibull*** | 141.5672 | 143.2336 | 0.00054 | 3.9194 | 12.12045 | 21.88648 | 29.19101 | 297.62944 | |

| | LOS CDI POSITIVE - slow diagnosis | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | **Min** | **1st Quantile** | **Median** | **Mean** | **3rd Quantile** | **Max** | **N observations above 100 days** | **N observations above 200 days** | **Comments** |
| **Original summary data** | | | **19.5** | **30.3** | | | | | |
| Exponential | 0.0005 | 8.7674 | 21.0485 | 30.3742 | 42.0335 | 352.5776 | 3791 | 145 | Second best distribution - lowest maximum value, fits also better with median |
| Lognormal | 0.4078 | 10.2089 | 19.3513 | 30.0956 | 36.5003 | 932.616 | 4040 | 664 | |
| ***Weibull*** | 0 | 6.156 | 17.781 | 30.3 | 40.794 | 456.528 | 5318 | 522 | |
| | **LOS CDI POSITIVE - average diagnosis** | | | | | | | | |
| | **Min** | **1st Quantile** | **Median** | **Mean** | **3rd Quantile** | **Max** | **N observations above 100 days** | **N observations above 200 days** | **Comments** |
| **Original summary data** | | | **20** | **26.9** | | | | | |
| Exponential | 0.00003 | 7.68333 | 18.61909 | 26.9345 | 37.59698 | 311.5671 | 2440 | 65 | Second best distribution - lowest maximum value, fits also better with median |
| Lognormal | 0.4627 | 10.1534 | 18.3612 | 26.9409 | 32.9434 | 650.1521 | 2635 | 318 | |
| ***Weibull*** | 0.0003 | 6.7493 | 17.4748 | 26.7551 | 36.6819 | 348.1296 | 2993 | 129 | |
| | **LOS CDI POSITIVE - rapid diagnosis** | | | | | | | | |

| | Min | 1st Quantile | Median | Mean | 3rd Quantile | Max | N observations above 100 days | N observations above 200 days | Comments |
|---|---|---|---|---|---|---|---|---|---|
| **Original summary data** | | | **20** | **23.2** | | | | | |
| Exponential | 0.00017 | 6.67094 | 16.13568 | 23.23045 | 32.282 | 307.7539 | 1293 | 13 | Second best distribution - lowest maximum value, fits also better with median |
| Lognormal | 0.3063 | 8.5956 | 15.6251 | 23.2293 | 28.4629 | 1056.037 | 1845 | 225 | |
| ***Weibull*** | 0.00006 | 5.68658 | 14.87688 | 23.17591 | 31.60179 | 298.5881 | 1946 | 661 | |

**Table N-5 Summary parameters for time-to-event variables**

| Parameter | Alpha (shape parameter) | Beta (scale parameter) | Distribution |
|---|---|---|---|
| **Duration symptoms CDI negative** | 0.5543052 | 4.23548 | Weibull |
| **Duration symptoms CDI positive** | 0.5880247 | 7.47795 | Weibull |
| **LOS CDI negative** | 0.8793202 | 26.57436 | Weibull |
| **LOS CDI positive – slow time to diagnosis** | 0.8389166 | 27.61969 | Weibull |
| **LOS CDI positive – average time to diagnosis** | 0.9315079 | 26.0319 | Weibull |
| **LOS CDI positive – rapid time to diagnosis** | 0.9145084 | 22.23937 | Weibull |

# Appendix O

# Model analysis and results

This appendix provides additional information on the model analysis implementation and additional results from the model described in Chapter 5 and Chapter 6.

## O.1 Model analysis methods

### O.1.1 Model initialisation

**Figure O-1 Assessment of how long the warm-up period should be to stabilise the number of new secondary cases in general ward – HT testing strategy at 15 minutes turnaround time**

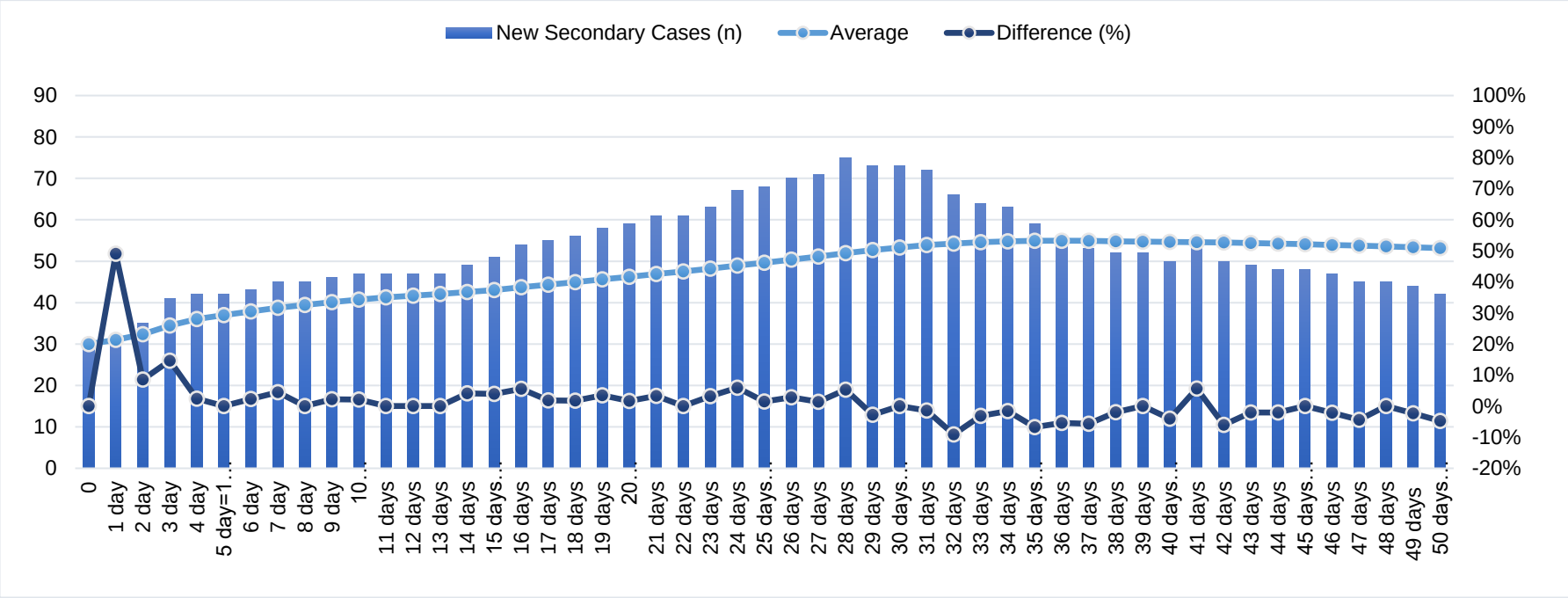
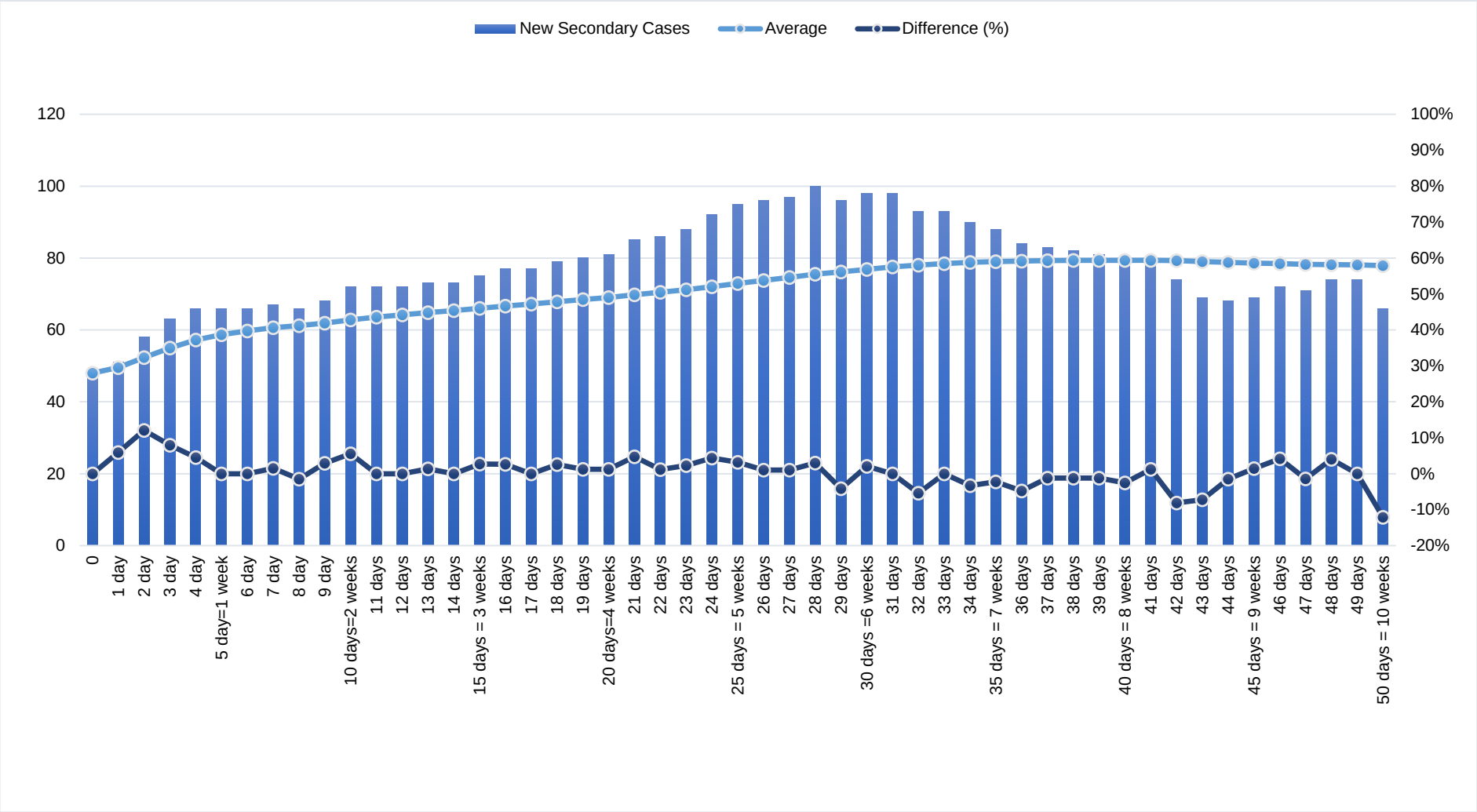


**Figure O-2 Assessment of how long the warm-up period should be to stabilise the number of new secondary cases in general ward – LTHT testing strategy**



## Model replication estimation

**Figure O-3 Assessment of how many model replications are required to stabilise the number of new secondary cases in the general ward – HT testing strategy at 15-minutes turnaround time**

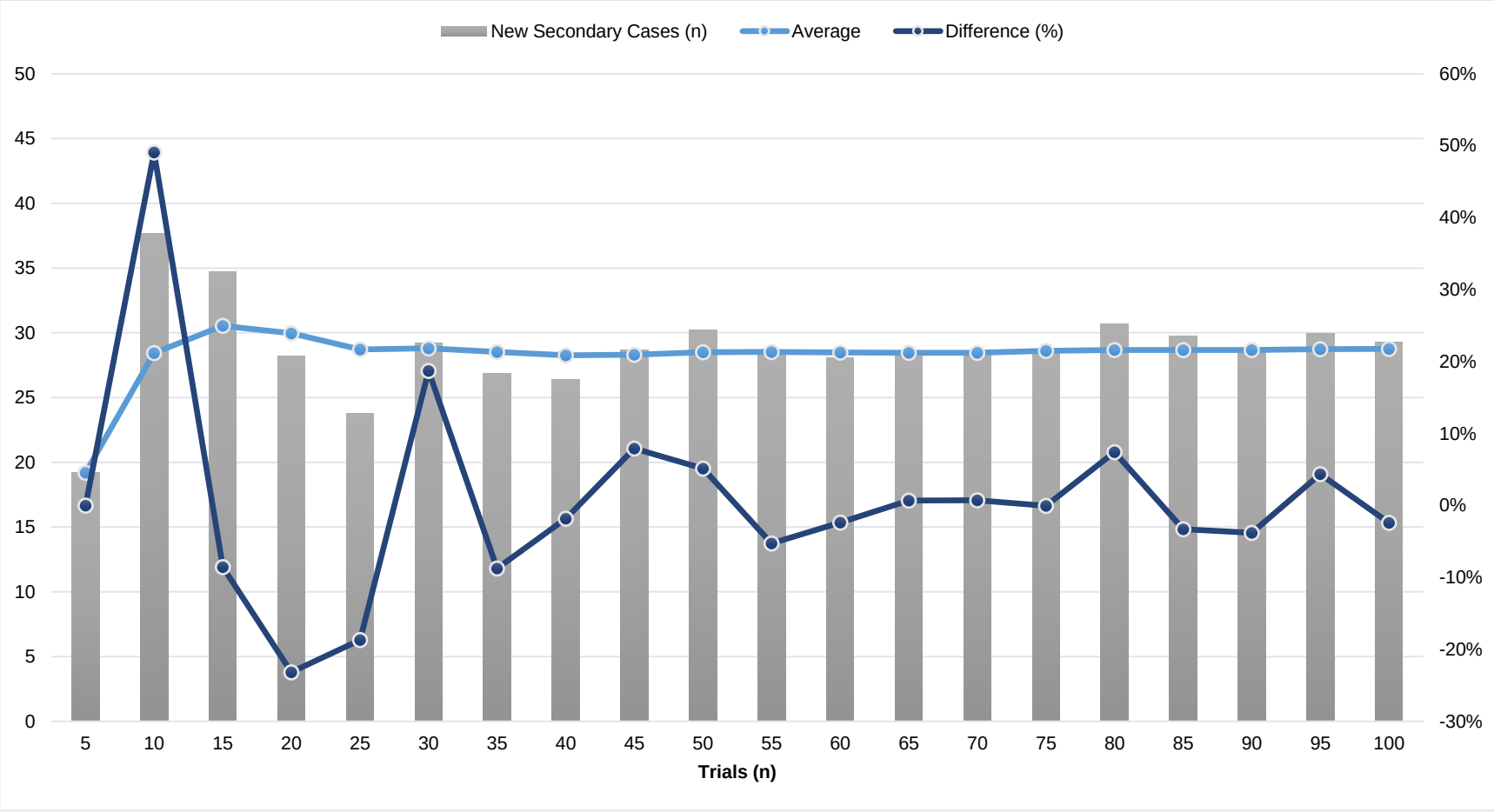
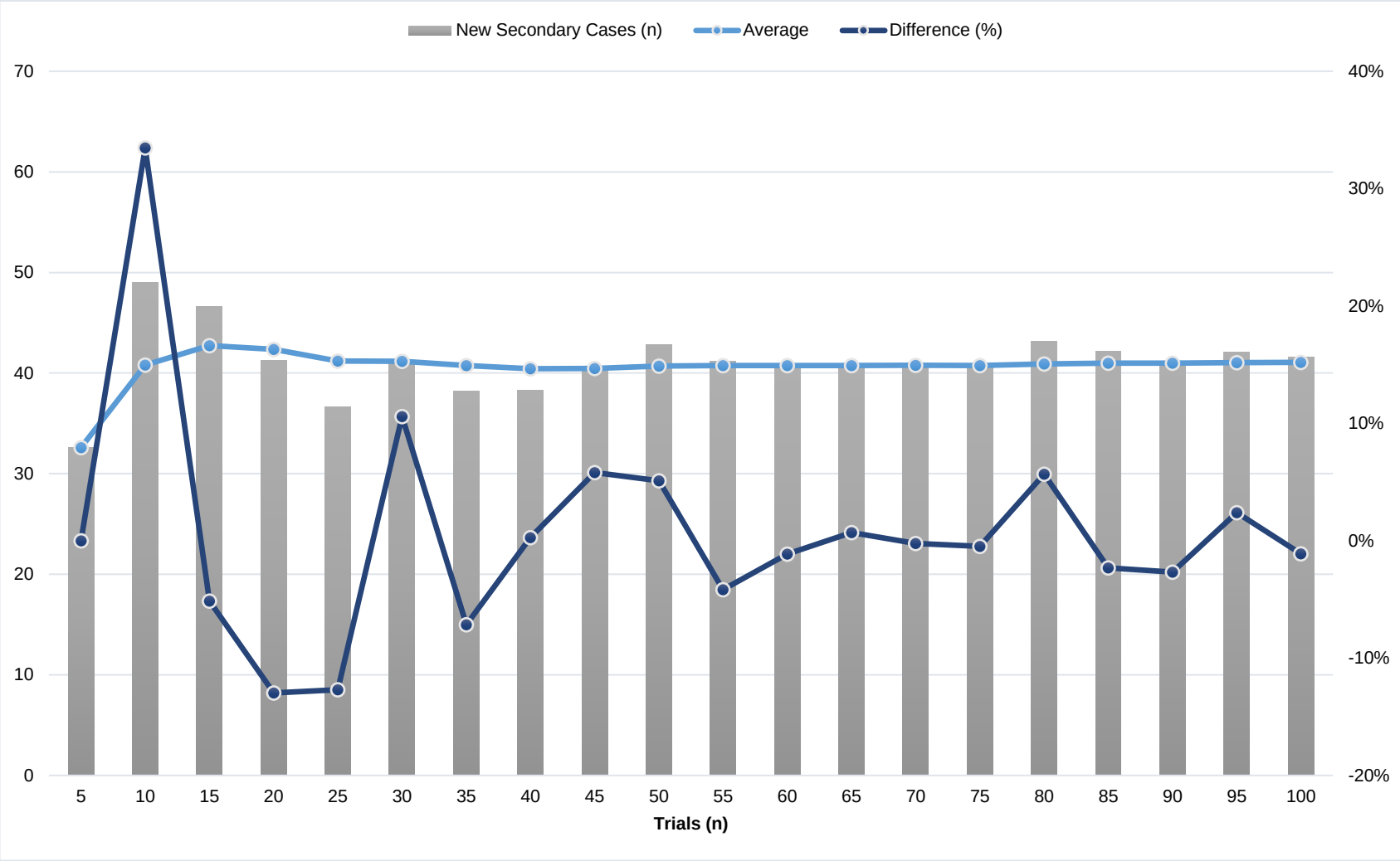


**Figure O-4 Assessment of how many model replications required to stabilise the number of new secondary cases in the general ward – LTHT testing strategy**



### O.1.2 Identification of minimum performance requirements

**Phase 2: Identification of key model drivers**

Upper and lower values were informed by literature estimates if available. If not available, a +/- 25% and +/- 50% deviation was applied.

For parameters where IPD were available, a decrease or increase of 25% and 50% from baseline IPD values was applied and then the scale and shape parameters for the fitted Weibull distributions were re-calculated to estimate the lower and upper bound.

For Weibull parameters based on aggregate data from the Barbut, F. et al. (260) study, the scale and shape parameters were re-calculated by applying a +/- 25% and 50% change to the mean estimates, holding the standard deviation constant.

**Table O-1 Lower and upper bounds for each model parameter, alongside calculations and sources, used to estimate the ranges for the univariate sensitivity analysis based on literature estimates and expert opinion**

| Parameter | Lower bound | Upper bound | Sources and calculations |
|---|---|---|---|
| **Reproductive ratio for CDI** | 0.55 | 1.99 | Lanzas, C. et al. (293) |
| **GDH EIA diagnostic sensitivity** | 94% | 96% | Crobach, M.J.T. et al. (222) |
| **GDH EIA diagnostic specificity** | 94% | 95% | Crobach, M.J.T. et al. (222) |
| **PCR diagnostic sensitivity** | 92% | 97% | Crobach, M.J.T. et al. (222) |
| **PCR diagnostic specificity** | 97% | 99% | Crobach, M.J.T. et al. (222) |
| **CCNA diagnostic sensitivity** | 75.8% | 100% | Planche, T. and M. Wilcox (294) – lower bound based on Barbut, F. et al. (311) estimate; upper bound set equal to perfect diagnostic sensitivity |

| | | | |
|---|---|---|---|
| **CCNA diagnostic specificity** | 96.7% | 100% | Planche, T. and M. Wilcox (294) – lower bound based on DiPersio, J.R. *et al.* (373) estimate; upper bound set equal to perfect diagnostic specificity |
| **Time-to-obtain stool sample** | 0 day | 2 days | Jones, W.S. et al. (266) |
| **Time-to-transport sample to the laboratory** | 0 minutes | 300 minutes | Assumption – high value is half-a-day as it requires transporting samples to an off-site laboratory |
| **GDH EIA time-to-result** | 45 minutes | 180 minutes | Assumption |
| **PCR time-to-result** | 30 minutes | 120 minutes | Assumption |
| **CCNA time-to-result** | 300 minutes | 3 days | Assumption |
| **Multiplex GI panel operating time** | 30 minutes | 50 minutes | Assumption |

**Table O-2 Lower and upper bounds for each model parameter, alongside calculations and sources, used to estimate the ranges for the univariate sensitivity analysis based arbitrary +/-25% and 50% deviation from baseline value**

| | +/- 25% deviation | | +/- 50% deviation | | Sources and calculations |
|---|---|---|---|---|---|
| **Parameter** | **Lower bound** | **Upper bound** | **Lower bound** | **Upper bound** | |
| **Disease prevalence CDI** | 0% | 25% | 0% | 50% | Arbitrary +/-25% and 50% deviation from baseline value |
| **Disease prevalence other GI pathogens** | 0% | 25% | 0% | 50% | Arbitrary +/-25% and 50% deviation from baseline value |

| | | | | | |
|---|---|---|---|---|---|
| **LOS CDI negative patients (days)** | Shape: 0.8790504 Scale: 19.9177571 | Shape: 0.8790134 Scale: 33.2105691 | Shape: 0.9194015 Scale: 13.9188038 | Shape: 0.8944039 Scale: 40.5561513 | COMBACTE-CDI CRF dataset – uplift or lower IPD baseline value by 25% and 50% and re-calculate shape and scale parameters for Weibull distribution |
| **LOS CDI positive patients – slow diagnosis (days)** | Shape: 0.6491104 Scale: 16.60751 | Shape: 1.043657 Scale: 38.52885 | Shape: 0.4774069 Scale: 6.952212 | Shape: 1.261911 Scale: 48.95625 | Barbut, F. et al. (260)– uplift or lower baseline value by 25% and 50% and re-calculate shape and scale parameters for Weibull distribution |
| **LOS CDI positive patients – average diagnosis (days)** | Shape: 0.7128001 Scale: 16.2108 | Shape: 1.167139 Scale: 35.4809 | Shape: 0.5159793 Scale: 7.141579 | Shape: 1.417708 Scale: 44.4119 | Barbut, F. et al. (260) – uplift or lower baseline value by 25% and 50% and re-calculate shape and scale parameters for Weibull distribution |
| **LOS CDI positive patients – rapid diagnosis (days)** | Shape: 0.7011098 Scale: 13.76674 | Shape: 1.144493 Scale: 30.42887 | Shape: 0.5075872 Scale: 5.961175 | Shape: 1.387367 Scale: 38.12686 | Barbut, F. et al. (260) – uplift or lower baseline value by 25% and 50% and re-calculate shape and scale parameters for Weibull distribution |
| **Duration diarrhoea after sampling – CDI negative patients (days)** | Shape: 0.5542175 Scale: 3.175754 | Shape: 0.55433 Scale: 5.294004 | Shape: 0.6305055 Scale: 2.7872972 | Shape: 0.5571209 Scale: 6.7335746 | COMBACTE-CDI CRF dataset – uplift or lower IPD baseline value by 25% and 50% and re-calculate shape and scale parameters for Weibull distribution |

| | | | | | |
|---|---|---|---|---|---|
| **Duration diarrhoea after sampling – CDI positive patients (days)** | Shape: 0.5879807 Scale: 5.610213 | Shape: 0.5879725 Scale: 9.344171 | Shape: 0.6373938 Scale: 4.3010929 | Shape: 0.594098 Scale: 11.598998 | COMBACTE-CDI CRF dataset – uplift or lower IPD baseline value by 25% and 50% Shand re-calculate shape and scale parameters for Weibull distribution |
| **Utility weight for adult UK inpatient with the first episode of CDI** | 32% | 52% | 21% | 63% | Arbitrary +/-25% and 50% deviation from baseline value |
| **Decrement utility weight due to inappropriate antibiotics** | 8% | 13% | 5% | 15% | Arbitrary +/-25% and 50% deviation from baseline value |
| **Probability of clinical cure – slow diagnosis** | 63.975% | 100% | 42.65% | 100% | Barbut, F. et al. (260) – decrease baseline value by 25% and 50% to estimate lower bound; set upper bound equal to 100% |
| **Probability of clinical cure – average diagnosis** | 68.025% | 100% | 45.35% | 100% | Barbut, F. et al. (260) – decrease baseline value by 25% and 50% to estimate lower bound; set upper bound equal to 100% |
| **Probability of clinical cure – rapid diagnosis** | 71.7% | 100% | 47.8% | 100% | Barbut, F. et al. (260) – decrease baseline value by 25% and 50% to estimate lower bound; set upper bound equal to 100% |
| **Cost of bed day in adult isolation** | £519.6225 | £866.0375 | £346.415 | £1039.245 | Arbitrary +/-25% and 50% deviation from baseline value |
| **Cost of bed day in general ward** | £437.25 | £728.75 | £291.5 | £874.5 | Arbitrary +/-25% and 50% deviation from baseline value |
| **GDH EIA cost per kit** | £3.63 | £6.05 | £2.42 | £7.26 | Arbitrary +/-25% and 50% deviation from baseline value |

| | | | | | |
|---|---|---|---|---|---|
| **PCR cost per run** | £20.175 | £33.625 | £13.45 | £40.35 | Arbitrary +/-25% and 50% deviation from baseline value |
| **CCNA cost** | £2.805 | £4.675 | £1.87 | £5.61 | Arbitrary +/-25% and 50% deviation from baseline value |
| **Multiplex GI panel cost per sample** | £32.2725 | £53.7875 | £21.515 | £64.545 | Arbitrary +/-25% and 50% deviation from baseline value |
| **Additional cost per secondary case per day** | £717.885 | £1196.475 | £478.59 | £1435.77 | Arbitrary +/-25% and 50% deviation from baseline value |
| **Vancomycin cost** | £99.3675 | £165.6125 | £66.245 | £198.735 | Arbitrary +/-25% and 50% deviation from baseline value |

**Table O-3 Summary of the key structural assumptions underpinning UK Teaching Hospitals and UK District Hospitals as simulated in the baseline scenario and S.4 scenario, respectively**

| Feature of clinical setting | UK Teaching Hospital | UK District Hospital | Calculations | Source |
|---|---|---|---|---|
| **Single rooms being available (n)** | 279 | 192 | Assumption – one-third of the reported single rooms in the COMBACTE CDI survey would be allocated for patients suspected with diarrhoea. See section 5.3.5.4 in Chapter 5 | COMBACTE-CDI survey |
| **Single rooms simulated (n)** | 93 | 64 | | |
| **Monthly average of stools being tested – median (n)** | 1430 | 850 | See section 5.3.5.5 in Chapter 5 | COMBACTE-CDI survey |
| **Laboratory** | On-site (15 minutes required to transport samples from ward to the laboratory) | Off-site ( two shifts: one in the morning and one at lunch time) | NA | Expert opinion |

**Table O-4 Second-best fitting distribution and related parameters for each of the time-to-event parameters**

| Time-to-event parameter | Distribution used at baseline | Second-best fitting distribution | Parameters for second-best fitting distributions |
|---|---|---|---|
| **S.5 (A) – LOS patients negative to CDI** | Weibull | Exponential | Rate = 0.03496115<br>Mean = 28.6031781 |
| **S.5 (B) – LOS patients positive to CDI (slow time-to-diagnosis)** | Weibull | Exponential | Rate = 0.0330033<br>Mean = 30.3000003 |
| **S.5 (C) – LOS patients positive to CDI (average time-to-diagnosis)** | Weibull | Exponential | Rate = 0.03717472<br>Mean = 26.90000086 |
| **S.5 (D) – LOS patients positive to CDI (rapid time-to-diagnosis)** | Weibull | Exponential | Rate = 0.04310345<br>Mean = 23.19999907 |
| **S.5 (E) – Duration of symptoms patients negative to CDI** | Weibull | Gamma | Alpha = 0.41780903<br>Beta = 0.05934788 |
| **S.5 (F) – Duration of symptoms patients positive to CDI** | Weibull | Gamma | Alpha = 0.4486443<br>Beta = 0.0365449 |

**Table O-5 Details on the additional labour costs for each test part of the standard care strategy applied to scenario 6**

| Testing strategy | Additional labour costs | Total costing estimate (including labour costs) | Source | Notes |
|---|---|---|---|---|
| **GDH EIA** | £5.66 | £10.5 | (267) | Original estimate was inflated from 2013 to 2021 |
| **PCR** | £1.19 | £28.09 | (300) | Original estimate was inflated from 2011 to 2021 |
| **CCNA** | £4.78 | £8.52 | (300) | Original estimate was inflated from 2011 to 2021 |
| **Multiplex GI panel** | £1.19 | £44.22 | (300) | Original estimate was inflated from 2011 to 2021 |

## O.2 Results

### Phase 2: Deterministic sensitivity analysis

For analysis purposes, parameters were divided into the following clusters:

- **disease-related parameters** – parameters relating to patient characteristics (e.g. prevalence of CDI and other GI pathogens), duration of symptoms, hospital LOS, treatment effectiveness at day 10 of antibiotic treatment, reproductive ratio, and health-related utility weights.
- **testing workflow parameters** – parameters relating to diagnostic accuracy and time-to-diagnosis (e.g. time-to-obtain stool samples, time-to-transport samples to the laboratory, operating time for each testing option).
- **costs**

The following sections report the results for the deterministic sensitivity analyses using estimates from the literature, and the +/- 25% and +/- 50% deviations from the baseline value, separately.

## *Phase 2 deterministic sensitivity analysis: +/- 25% deviation from the baseline value*

### Table O-6 Secondary cases prevented, incremental QALY gains, incremental costs INMB and INMB range for HT over LTHT testing strategies compared to baseline results – sorted by lower and upper bound for each disease-related parameter using estimates from the literature and the +/- 25% deviation from the baseline value

| Parameter | Bound | Secondary cases prevented | Incremental QALY gains | Incremental costs | INMB at £20,000 per QALY gained | INMB Range |
|---|---|---|---|---|---|---|
| **Baseline results** | NA | 3.00 | 0.02 | £410 | £0 | NA |
| **Disease prevalence CDI** | Upper | -13.10 | 0.57 | -£734,620 | £746,064 | £885,093 |
| | Lower | 0.00 | -0.13 | £136,412 | -£139,029 | |
| **Disease prevalence other GI pathogens** | Upper | 1.21 | 0.05 | -£82,319 | £83,343 | £110,077 |
| | Lower | 3.24 | 0.02 | £27,123 | -£26,734 | |
| **LOS CDI negative patients** | Upper | 4.74 | 0.00 | £40,459 | -£40,526 | £96,233 |
| | Lower | 0.61 | 0.06 | -£54,569 | £55,707 | |
| **LOS CDI positive patients - slow diagnosis** | Upper | -5.19 | 0.17 | -£501,234 | £504,583 | £959,857 |
| | Lower | 10.30 | -0.10 | £453,357 | -£455,274 | |
| **LOS CDI positive patients - average diagnosis** | Upper | 3.41 | 0.02 | £32,424 | -£32,121 | £39,815 |
| | Lower | 3.21 | 0.02 | -£7,325 | £7,694 | |
| **LOS CDI positive patients - rapid diagnosis** | Upper | 7.74 | -0.05 | £348,284 | -£349,372 | £666,371 |
| | Lower | -1.40 | 0.09 | -£315,210 | £317,000 | |
| **Duration diarrhoea CDI negative patients** | Upper | 3.06 | 0.023 | £1,574 | -£1,118 | £201,530 |
| | Lower | -0.51 | 0.078 | -£198,857 | £200,412 | |
| **Duration diarrhoea CDI positive patients** | Upper | 3.00 | 0.02 | £394 | £16 | £54 |
| | Lower | 3.00 | 0.02 | £449 | -£38 | |
| **Probability of clinical cure - slow diagnosis** | Upper | 3.06 | -0.02 | £10,659 | -£11,079 | £59,775 |
| | Lower | 2.07 | 0.11 | -£46,581 | £48,696 | |
| **Probability of clinical cure - average diagnosis** | Upper | 3.00 | 0.02 | -£1,913 | £2,378 | £4,593 |
| | Lower | 3.00 | 0.02 | £2,550 | -£2,215 | |
| **Probability of clinical cure - rapid diagnosis** | Upper | 2.91 | 0.03 | -£3,597 | £4,233 | £55,856 |
| | Lower | 3.89 | -0.07 | £50,182 | -£51,623 | |
| **Reproductive ratio for CDI** | Upper | 5.6 | -0.02 | £78,358 | -£78,779 | £120,646 |
| | Lower | 1.5 | 0.04 | -£41,016 | £41,867 | |
| **Utility weight for adult CDI patient** | Upper | 3.0 | -0.01 | £410 | -£587 | £1,170 |
| | Lower | 3.0 | 0.05 | £410 | £583 | |
| **Decrement utility weight due to inappropriate antibiotics** | Upper | 3.0 | -0.01 | £410 | -£587 | £1,170 |
| | Lower | 3.0 | 0.05 | £410 | £583 | |

For each model parameter value, the number of secondary cases being prevented, incremental QALY gains, incremental costs and INMB is compared against the respective result at baseline. A green up arrow is associated with model parameter values which lead to model outcomes to be more clinically- and/or cost-effective than the respective results at baseline. A yellow bar is associated with model parameter values which lead to model outcomes equal to the respective results at baseline. A red down arrow is associated with model parameter values which lead to the model results to be less clinically- and/or cost-effective than the respective results at baseline.

**Table O-7 Secondary cases prevented, incremental QALY gains, incremental costs INMB and INMB range for HT over LTHT testing strategies compared to baseline results – sorted by lower and upper bound for each testing workflow-related parameter using estimates from the literature and +/- 25% deviation from the baseline value**

| Parameter | Bound | Secondary cases prevented (n) | Incremental QALY gains | Incremental costs | INMB at £20,000 per QALY gained | Range |
|---|---|---|---|---|---|---|
| **Baseline results** | NA | 3 | 0.0205 | £410 | £0 | NA |
| **GDH EIA sensitivity** | Upper | 3.5 | -0.02916 | £12,461 | -£13,045 | £34,806 |
| | Lower | 2.3 | 0.044 | -£20,882 | £21,762 | |
| **GDH EIA specificity** | Upper | 3.2 | 0.01556 | £10,411 | -£10,100 | £10,100 |
| | Lower | 2.9 | 0.02051 | £410 | £0 | |
| **PCR sensitivity** | Upper | 3.1 | 0.01413 | £4,688 | -£4,406 | £13,800 |
| | Lower | 2.7 | 0.03725 | -£8,650 | £9,395 | |
| **PCR specificity** | Upper | 3.1 | 0.01413 | £4,688 | -£4,406 | £25,276 |
| | Lower | 2.5 | 0.03425 | -£20,186 | £20,871 | |
| **CCNA sensitivity** | Upper | 3.5 | 0.0013 | £16,597 | -£16,571 | £30,444 |
| | Lower | 2.5 | 0.03732 | -£13,127 | £13,873 | |
| **CCNA specificity** | Upper | 3.1 | 0.01627 | £5,810 | -£5,485 | £32,714 |
| | Lower | 2.2 | 0.03913 | -£26,446 | £27,229 | |
| **Time to obtain stool samples** | Upper | 3.8 | 0.01812 | £10,705 | -£10,343 | £418,222 |
| | Lower | -8.9 | 0.21715 | -£403,536 | £407,879 | |
| **Time to transport sample to the laboratory** | Upper | 8.0 | -0.08884 | £283,336 | -£285,113 | £316,336 |
| | Lower | 2.5 | 0.03715 | -£30,480 | £31,223 | |
| **GDH EIA operating time** | Upper | 0.2 | 0.08451 | -£87,243 | £88,933 | £89,135 |
| | Lower | 3.0 | 0.01905 | £583 | -£202 | |
| **PCR operating time** | Upper | 3.0 | 0.02051 | £464 | -£53 | £53 |
| | Lower | 3.0 | 0.02051 | £410 | £0 | |
| **CCNA operating time** | Upper | 2.9 | 0.11223 | -£43,427 | £45,671 | £70,218 |
| | Lower | 3.8 | -0.02011 | £24,144 | -£24,547 | |
| **Multiplex GI panel operating time** | Upper | 3.5 | -0.62396 | £1,080,617 | -£1,093,096 | £1,348,716 |
| | Lower | -2.7 | 0.15784 | -£252,463 | £255,620 | |

For each model parameter value, the number of secondary cases being prevented, incremental QALY gains, incremental costs and INMB is compared against the respective result at baseline. A green up arrow is associated with model parameter values which lead to model outcomes to be more clinically- and/or cost-effective than the respective results at baseline. A yellow bar is associated with model parameter values which lead to model outcomes equal to the respective results at baseline. A red down arrow is associated with model parameter values which lead to the model results to be less clinically- and/or cost-effective than the respective results at baseline.

**Table O-8 Secondary cases prevented, incremental QALY gains, incremental costs INMB and INMB range for HT over LTHT testing strategies compared to baseline results – sorted by lower and upper bound for each cost parameter using estimates from the literature and the +/- 25% deviation from the baseline value**

| Parameter | Bound | Secondary cases prevented (n) | Incremental QALY gains | Incremental costs | INMB at £20,000 per QALY gained | INMB at £20,000 per QALY gained |
|---|---|---|---|---|---|---|
| **Baseline results** | NA | 3.00 | 0.02 | £410 | £0 | NA |
| **Cost of bed day in adult isolation** | Upper | 3.0 | 0.02 | £42,149 | -£41,739 | £82,696 |
| | Lower | 3.0 | 0.02 | -£40,547 | £40,957 | |
| **Cost of bed day in general ward** | Upper | 3.0 | 0.02 | -£55,710 | £56,120 | £112,238 |
| | Lower | 3.0 | 0.02 | £56,529 | -£56,118 | |
| **GDH EIA cost per kit** | Upper | 3.0 | 0.02 | £316 | £95 | £189 |
| | Lower | 3.0 | 0.02 | £505 | -£95 | |
| **PCR cost per run** | Upper | 3.0 | 0.02 | -£71 | £481 | £962 |
| | Lower | 3.0 | 0.02 | £891 | -£481 | |
| **CCNA cost** | Upper | 3.0 | 0.02 | £259 | £151 | £303 |
| | Lower | 3.0 | 0.02 | £561 | -£151 | |
| **Multiplex GI panel cost per sample** | Upper | 3.0 | 0.02 | £410 | £0 | £0 |
| | Lower | 3.0 | 0.02 | £410 | £0 | |
| **Additional cost per secondary case per day** | Upper | 3.0 | 0.02 | £21,331 | -£20,921 | £41,842 |
| | Lower | 3.0 | 0.02 | -£20,511 | £20,921 | |
| **Vancomycin cost** | Upper | 3.0 | 0.02 | £3,263 | -£2,853 | £5,341 |
| | Lower | 3.0 | 0.02 | -£2,078 | £2,488 | |

For each model parameter value, the number of secondary cases being prevented, incremental QALY gains, incremental costs and INMB is compared against the respective result at baseline. A green up arrow is associated with model parameter values which lead to model outcomes to be more clinically- and/or cost-effective than the respective results at baseline. A yellow bar is associated with model parameter values which lead to model outcomes equal to the respective results at baseline. A red down arrow is associated with model parameter values which lead to the model results to be less clinically- and/or cost-effective than the respective results at baseline.

***Phase 2 deterministic sensitivity analysis: +/- 50% deviation from the baseline value***

**Figure O-5 Tornado plot showing the results of the deterministic univariate sensitivity analysis for the most influential model parameters using estimates from the literature and the +/- 50% deviation from baseline values. Results are expressed in terms of INMB at £20,0000 per QALY gained. INMB at baseline is equal to £0 and reflects the cost-effectiveness output associated with the minimum performance specifications for HT at 15-minutes turnaround time**

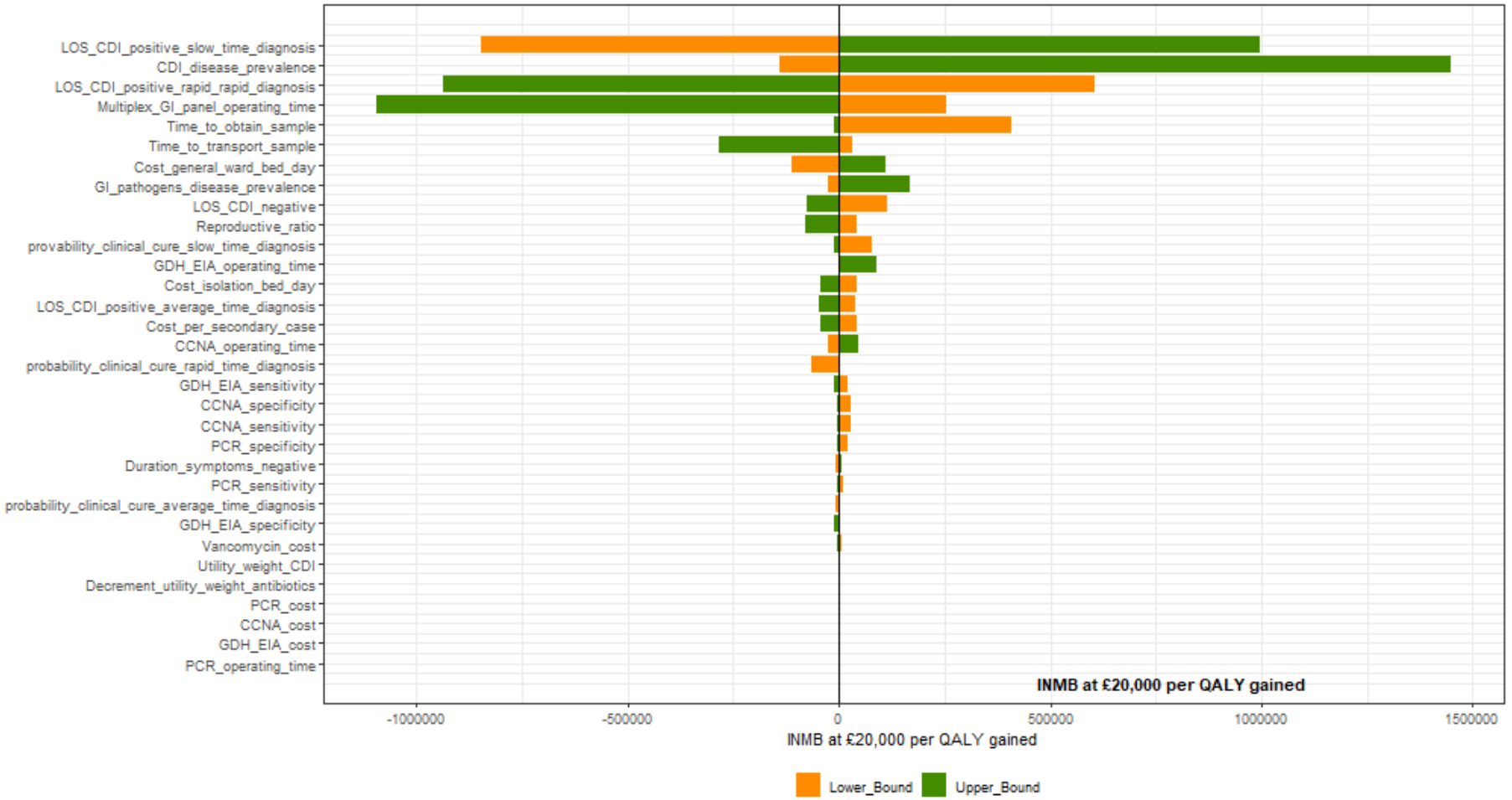


**Table O-9 Secondary cases prevented, incremental QALY gains, incremental costs INMB and INMB range for HT over LTHT testing strategies compared to baseline results – sorted by lower and upper bound for each disease-related parameter using estimates from the literature and the +/- 50% deviation from the baseline value**

| Parameter | Bound | Secondary cases prevented | Incremental QALY gains | Incremental costs | INMB at £20,000 per QALY gained | INMB Range |
|---|---|---|---|---|---|---|
| **Baseline results** | NA | 3.00 | 0.02 | £410 | £0 | NA |
| **Disease prevalence CDI** | Upper | -30.46 | 1.34 | -£1,419,974 | £1,446,710 | £1,585,739 |
| | Lower | 0.00 | -0.13 | £136,412 | -£139,029 | |
| **Disease prevalence other GI pathogens** | Upper | -0.41 | 0.07 | -£164,944 | £166,414 | £193,148 |
| | Lower | 3.24 | 0.02 | £27,123 | -£26,734 | |
| **LOS CDI negative patients** | Upper | 6.37 | -0.03 | £76,553 | -£77,124 | £192,325 |
| | Lower | -1.96 | 0.10 | -£113,239 | £115,201 | |
| **LOS CDI positive patients - slow diagnosis** | Upper | -12.77 | 0.30 | -£990,546 | £996,568 | £1,842,218 |
| | Lower | 16.53 | -0.20 | £841,648 | -£845,650 | |
| **LOS CDI positive patients - average diagnosis** | Upper | 3.44 | 0.01 | £47,612 | -£47,329 | £85,542 |
| | Lower | 2.70 | 0.03 | -£37,698 | £38,213 | |
| **LOS CDI positive patients - rapid diagnosis** | Upper | 12.67 | -0.19 | £932,429 | -£936,178 | £1,540,127 |
| | Lower | -5.41 | 0.15 | -£600,899 | £603,949 | |
| **Duration diarrhoea CDI negative patients** | Upper | 3.00 | 0.027 | -£5,678 | £6,221 | £14,541 |
| | Lower | 3.21 | 0.020 | £8,722 | -£8,320 | |
| **Duration diarrhoea CDI positive patients** | Upper | 3.00 | 0.02 | £394 | £16 | £54 |
| | Lower | 3.00 | 0.02 | £449 | -£38 | |
| **Probability of clinical cure - slow diagnosis** | Upper | 3.06 | -0.02 | £10,659 | -£11,079 | £90,327 |
| | Lower | 1.59 | 0.18 | -£75,615 | £79,248 | |
| **Probability of clinical cure - average diagnosis** | Upper | 2.89 | 0.02 | -£1,913 | £2,378 | £11,290 |
| | Lower | 3.11 | 0.01 | £9,075 | -£8,912 | |
| **Probability of clinical cure - rapid diagnosis** | Upper | 2.91 | 0.03 | -£3,597 | £4,233 | £68,653 |
| | Lower | 3.71 | -0.14 | £61,579 | -£64,419 | |
| **Reproductive ratio for CDI** | Upper | 5.6 | -0.02 | £78,358 | -£78,779 | £120,646 |
| | Lower | 1.5 | 0.04 | -£41,016 | £41,867 | |
| **Utility weight for adult CDI patient** | Upper | 3.0 | -0.05 | £410 | -£1,444 | £2,887 |
| | Lower | 3.0 | 0.09 | £410 | £1,444 | |
| **Decrement utility weight due to inappropriate** | Upper | 3.0 | -0.04 | £410 | -£1,166 | £2,333 |
| | Lower | 3.0 | 0.08 | £410 | £1,166 | |

costs and INMB is compared against the respective result at baseline. A green up arrow is associated with model parameter values which lead to model outcomes to be more clinically- and/or cost-effective than the respective results at baseline. A yellow bar is associated with model parameter values which lead to model outcomes equal to the respective results at baseline. A red down arrow is associated with model parameter values which lead to the model results to be less

**Table O-10 Secondary cases prevented, incremental QALY gains, incremental costs INMB and INMB range for HT over LTHT testing strategies compared to baseline results – sorted by lower and upper bound for each cost parameter using estimates from the literature and the +/- 50% deviation from the baseline value**

| Parameter | Bound | Secondary cases prevented (n) | Incremental QALY gains | Incremental costs | INMB at £20,000 per QALY gained | INMB at £20,000 per QALY gained |
|---|---|---|---|---|---|---|
| **Baseline results** | NA | 3.00 | 0.02 | £410 | £0 | NA |
| **Cost of bed day in adult isolation** | Upper | 3.0 | 0.02 | £43,463 | -£43,053 | £86,108 |
| | Lower | 3.0 | 0.02 | -£42,645 | £43,055 | |
| **Cost of bed day in general ward** | Upper | 3.0 | 0.02 | -£111,827 | £112,238 | £224,471 |
| | Lower | 3.0 | 0.02 | £112,644 | -£112,233 | |
| **GDH EIA cost per kit** | Upper | 3.0 | 0.02 | £221 | £189 | £378 |
| | Lower | 3.0 | 0.02 | £599 | -£189 | |
| **PCR cost per run** | Upper | 3.0 | 0.02 | -£552 | £962 | £1,925 |
| | Lower | 3.0 | 0.02 | £1,373 | -£962 | |
| **CCNA cost** | Upper | 3.0 | 0.02 | £108 | £303 | £605 |
| | Lower | 3.0 | 0.02 | £713 | -£303 | |
| **Multiplex GI panel cost per sample** | Upper | 3.0 | 0.02 | £410 | £0 | £0 |
| | Lower | 3.0 | 0.02 | £410 | £0 | |
| **Additional cost per secondary case per day** | Upper | 3.0 | 0.02 | £42,253 | -£41,842 | £83,685 |
| | Lower | 3.0 | 0.02 | -£41,432 | £41,842 | |
| **Vancomycin cost** | Upper | 3.0 | 0.02 | £5,386 | -£4,976 | £9,952 |
| | Lower | 3.0 | 0.02 | -£4,566 | £4,976 | |

For each model parameter value, the number of secondary cases being prevented, incremental QALY gains, incremental costs and INMB is compared against the respective result at baseline. A green up arrow is associated with model parameter values which lead to model outcomes to be more clinically- and/or cost-effective than the respective results at baseline. A yellow bar is associated with model parameter values which lead to model outcomes equal to the respective results at baseline. A red down arrow is associated with model parameter values which lead to the model results to be less clinically- and/or cost-effective than the respective results at baseline.

## O.3 Expected return of investment

**Table O-11 Calculations of expected return of investment for a 15-minute rapid test for CDI using a threshold price equal to £44**

| | UK District Hospital | UK Teaching Hospital | Source and notes |
|---|---|---|---|
| Monthly average of stool samples tested (median) | 850 | 1430 | COMBACTE-CDI survey |
| Yearly average | 10200 | 17160 | COMBACTE-CDI survey |
| Proportion of samples tested for CDI only | 3060 | 5148 | COMBACTE-CDI survey - assuming 30% proportion |
| Total patients tested for CDI in the UK over a year (n) | 8208 | | |

| Portion of market | Expected return using threshold price at 15-minute turnaround time (£44) | 10 Years |
|---|---|---|
| 10% | £336115.2 | £361,152 |
| 20% | £72230.4 | £722,304 |
| 30% | £108345.6 | £1,083,456 |
| 40% | £144460.8 | £1,444,608 |