Test evaluation in a resource-limited country context from a health economics perspective

By

Samuel Nkansah Frempong

A thesis submitted to

The University of Birmingham

for the degree of

DOCTOR OF PHILOSOPHY

Health Economics Unit

Institute of Applied Health Research

College of Medical and Dental Sciences

The University of Birmingham

Birmingham, November 2018

Abstract

In Ghana, there are issues with the diagnosis of typhoid fever such as delays in diagnosis, concerns about the accuracy of current tests, and lack of availability. All of which highlight the need for the development of a rapid, accurate and easily accessible diagnostic test. Whilst several studies have indicated the importance of an iterative use of economic evaluations during the early phases of development of medical devices, there is little specific guidance on their implementation. The aim of this research was to examine the potential cost-effectiveness of a hypothetical rapid test for typhoid fever diagnosis in Ghana.

Two systematic reviews, a qualitative survey and an early cost-effectiveness analysis were conducted to achieve the research aim.

The results here show that there is a general lack of clarity on the methods used in the early economic evaluation of medical tests, and the techniques available to modellers that can demonstrate the value of conducting further research and product development (i.e., value of information (VOI) analysis, headroom analysis) should be better utilized. For a hypothetical test to perform better than the current available tests in terms of QALYs gained and cost-effectiveness, it is necessary that it has a high specificity (at least 70%) and should not be priced more than $10.

I would like to dedicate this thesis to my wonderful Family

Mrs Keren Frempong, Mama Vic, Mr Edward Appiah Frempong, Mr and Mrs Nkansa, Mr and Mrs Owusu-Ansah, Mark, Rowland, Bernice and Jesse

Acknowledgements

I am grateful to God for the opportunity and grace to complete this PhD. My sincere gratitude goes to my supervisors, Dr Pelham Barton, Dr Andrew John Sutton and Dr Clare Davenport for their guidance, assistance and wonderful support throughout this project. If I had to do it all over again, I would still have you as my supervisors.

I would like to thank the Government of Ghana for funding my PhD through the administration of The Ghana Education Trust fund.

I am also grateful to Dr Lavinia Ferrante di Ruffano for her insightful thoughts at the beginning of the research, Dr Vittal Mogasale of the International Vaccine Institute for his support and Mrs Susan Bayliss for her constructive suggestions in developing the search strategies for the systematic reviews.

I would also want to thank Dr Peter Hall and Dr Sue Jowett who graciously agreed to be my examiners.

I would also like to acknowledge all participants for their time, willingness to participate in the qualitative survey and their insightful responses.

Many thanks to Pastor Moses, Alfred Opukumo Wilson, Bibiane Manga Atangane, Adekemi Sekoni, Eniya Lufumpa, and Dr Raymond Oppong who made these 3 years worthwhile.

Declaration

The work presented in this thesis is the result of the original research conducted by the author Samuel Nkansah Frempong. During the period of the postgraduate study within the Health Economics Unit, University of Birmingham, the following articles were accepted for publication. The secondary authors advised on the study design, data analysis and paper editing.

Chapter 3

Frempong, S.N., Sutton, A.J., Davenport, C. and Barton, P., (2017): Economic evaluation of medical tests at the early phases of development: a systematic review of empirical studies. **Expert review of pharmacoeconomics & outcomes research**, 18(1), pp.13-23.

Chapter 4

Frempong, S.N., Davenport, C., Sutton, A.J., Nonvignon, J. and Barton, P., (2018): Integrating Qualitative Techniques in Model Development: A Case Study Using the Framework Approach. **Applied health economics and health policy**, 16(5), pp.723-733.

Chapter 5

Frempong, S.N., Sutton, A.J., Davenport, C. and Barton, P., (2018): Economic evaluation of typhoid–a review. **Expert review of pharmacoeconomics & outcomes research**, pp.1-7.

Table of Contents

LIST OF TABLES

LIST OF FIGURES

Chapter 1: Introduction

1.1 Background

Over the past few decades, the introduction of medical tests has accelerated substantially (Redekop, 2006; Koffijberg et al., 2013). Despite this increase, there exist challenges for the economic evaluation of tests which represent on-going research priorities (such as linking information from different studies that may not be comparable in the absence of test-treat randomised controlled trials) even in countries where research is relatively well developed (Merlin et al., 2013; Merlin, 2014). Resource-limited countries have additional challenges such as there being no proper systems in place to ensure the adequate evaluation of tests, and health informatics not being comprehensive or clearly linked to priority (Odame, 2013; Schroeder et al., 2015). Consequently, testing in some disease areas that pose serious public health threats and the resulting increased cost to health care systems are being neglected; one such disease is typhoid fever (Wain et al., 2015).

In Ghana (a resource-limited country), previously reported annual incidence of typhoid fever in 2017 was 60,892 cases (Kim et al., 2017) and this ranks among the leading 20 causes of outpatient illness in the country, accounting for 0.92% of all hospital admissions (Marks et al, 2010). The two tests for typhoid fever diagnosis in Ghana are the Widal test (a serological test) and blood culture (GNDP, 2014). Compared to blood culture, the Widal test is easier and requires fewer resources to perform, however, the Widal test is an inaccurate indicator of typhoid fever, resulting in inappropriate treatment frequently being administered in Ghana (GNDP, 2014). This has led to the development of resistance of *Salmonella typhi* and *paratyphi* (the causative organisms of typhoid fever) to previously effective cheap treatments such as chloramphenicol, and this has resulted in the current use of the more expensive ciprofloxacin

in Ghana (GNDP, 2014). Blood culture lacks sensitivity (the proportion of people with disease who will have a positive test result) (Wain et al., 2015), meaning that the disease will be missed in many patients (i.e., a high false negative rate) resulting in lost opportunities for treatment, which might lead to increased morbidity and mortality. Furthermore, culture requires equipment, supplies, trained laboratory personnel and electricity which may hardly be available in primary health care facilities in Ghana, and when performed, can take 2-3 days before test results are obtained. Thus, diagnosis may be delayed or overlooked because patients may not return for their lab results. Patients may end up with complications such as intestinal perforation resulting in death if treatment is postponed or patients are lost to follow up (Huang et al., 2013). Furthermore, those without typhoid fever may receive unnecessary and inappropriate antimicrobial treatment if treated presumptively (Keddy et al., 2011). Another issue is that typhoid fever clinically resembles other febrile conditions such as malaria and is easily misdiagnosed without laboratory confirmation which could lead to an overprescribing of antimalarial therapies associated with a huge economic impact on the Ghanaian economy. For example, Nonvignon at al. (2016) showed that businesses in Ghana lost about US$6.58 million to malaria in 2014, 90% of which was direct cost. Therefore, the development of a test for typhoid fever in the Ghanaian setting which is accurate, simple and affordable with a quick turnaround time to enable prompt implementation of an appropriate effective treatment for typhoid fever has an obvious attraction.

To comprehensively evaluate the potential value of a new test (which is the focus of this thesis) there is the need for an in-depth understanding of the principles underpinning test evaluation and the implementation of economic evaluations of medical tests (and for the purpose of this thesis, economic evaluations conducted at the early phases of test development). Economic evaluation is defined as “the comparative analysis of alternative courses of action in terms of

their cost and consequences" (Drummond et al., 2005, p.9). Central to economic evaluation is the fact that, an imbalance between available resources and the potential use for these resources will always exist (Morris et al., 2012). Therefore, choosing to deploy resource to one use means that the opportunity to use that resource in their next best alternative is foregone (i.e., there is an opportunity cost associated with every decision made) (Drummond et al., 2005). The purpose of economic evaluation is to ensure that there is good value for money spent. Economic evaluation is a tool that can help users make the most of resources, decide between promising interventions, and demonstrate the benefits of an intervention. Economic evaluations are either conducted alongside randomised controlled trials (RCTs) or using decision analytic modelling. In the case of medical tests, decision analytic modelling has been advocated as the most systematic and transparent method for assessing cost-effectiveness in the absence of test-treat RCTs (Trikalinos et al., 2009). Decision analytic modelling involves combining evidence from a number of sources to provide evidence on clinical and cost-effectiveness. There are are two main types of decision analytic models namely: static and dynamic models (Drummond et al., 2005). The evaluation of medical tests differs in some important ways from the evaluation of treatments (Kip et al., 2018). Medical testing affects the quality of life of a patient primarily in two ways. These are the direct effects associated with the test or the testing process (which could either be positive or negative), and the downstream benefits arising from clinical decisions made based on the test results (AHRQ, 2012). It is only by capturing these effects (which is dependent on an in-depth understanding of the care pathway) that the true value of a medical test can be established (Van den Bruel et al., 2007; Whiting et al., 2015). Knowledge of the care pathway is essential when evaluating the impact of testing on patient outcomes and should involve all stakeholders (those for whom the use of a test will have an impact personally or from an organisation perspective or from a population perspective). Economic evaluations are usually conducted at the late stage of product development (including medical tests)

(Pietzsch and Paté-Cornell, 2008), the rationale being that at this point there will be enough device-specific data for their proper evaluation (Redekop and Mikudina, 2013). However, economic evaluations conducted at the early phases of product development are conducted at a time when there is a lack of data (for example, lack of data describing the effectiveness of an intervention) and are more iterative in nature. The lack of data implies that economic evaluations conducted at the early phases of product development are associated with increased uncertainty (Buisman et al., 2016), and is likely to be more pertinent for tests. This is further compounded by the fact that, due to the indirect nature of tests in terms of their impact on patient relevant outcomes, data on test accuracy and downstream consequences may only be acquired by mapping clinical pathways (which is difficult when the role of a new test is not clear) leading to potentially more uncertain data. Furthermore, early economic evaluations are used to inform investment decisions, whilst late economic evaluations are used to inform adoption decisions.

Clearly, there exist some differences in the evaluation of medical tests compared to treatments, as well as in the implementation of economic evaluations at an early stage of technology development compared to the late stage. These differences highlight the need for a better understanding of the key principles of test evaluation needed to ensure the effective evaluation of the true value of a new test. Furthermore, these differences suggest that the methodologies used in the analysis of late economic evaluation need modification for use in early economic evaluation. An understanding of these principles is key to this study because the intervention being evaluated is a medical test which is at the early development phase, and there is little specific guidance on the implementation of economic evaluations at the early stages of technology development (Buisman et al., 2016).

1.2 Study aim and objectives

The aim of this research was to examine the potential cost-effectiveness of a hypothetical test (hereafter referred to as the HT-test) for typhoid fever diagnosis in Ghana. Specifically:

1. To examine the methodologies and tools that have been employed in the early economic evaluation of medical tests.
2. To explore the variations in existing test-treat strategies for typhoid fever in Ghana to delineate the care pathway for typhoid fever.
3. To explore the types of models that have been adopted for test-treat evaluations of typhoid fever and to capture data on model inputs that may be useful for a de novo model.
4. To design an economic model to estimate the maximum price and the minimum test performance required for the HT-test to be cost-effective in the Ghanaian setting.

1.3 Study design and relevance of study

This research was conducted over four distinct but interrelated stages (two systematic reviews, a qualitative survey, and early cost-effectiveness analysis). Each of these stages addressed one of the objectives outlined. The first systematic review addressed objective 1, the qualitative survey addressed objective 2, the second systematic review addressed objective 3 and the early cost-effectiveness analysis addressed objective 4. These are presented in Chapters 3, 4, 5 and 6 respectively.

The findings of this thesis provide valuable information to test developers about the market potential of developing a new test for typhoid fever that can be applied in the Ghanaian setting.

The findings also provide valuable information for decision makers to support their decision making on the further development of a rapid test for typhoid diagnosis in Ghana. This could help allocate limited budgets more efficiently and greatly improve the delivery of care for typhoid fever patients in Ghana. The findings also have wider implications for the early economic evaluation of tests by contributing to a limited evidence base and identifying areas requiring further research.

Decision maker in the context of adoption of health technologies in Ghana refers to the Government of Ghana through the management of the National Health Insurance Authority (NHIA). The NHIA are responsible for entering into purchasing agreements with providers and reimbursing those providers (Soddzi-Tettey et al., 2012). The decision criteria previously adopted by the NHIA in choosing which health interventions to fund and which not to fund were based on the World Health Organization's (WHO) suggested cost-effectiveness threshold based on multiples of a country's gross domestic product (GDP). That is, an intervention that per disability-adjusted life-year (DALY) avoided, costs less than three times the national annual GDP per capita is considered cost-effective, those that exceed this level are considered not cost–effective, whereas one that costs less than once the national annual GDP per capita is considered highly cost-effective. However, due to the recognition of major shortcomings associated with the use of this threshold, the WHO has recently withdrawn this threshold (Bertram et al., 2016) which has lead to the adoption of the new estimated country-specific threshold of less than $104-$951/QALY by the NHIA. This threshold has been calculated from estimates of opportunity cost, estimates of the relationship between the Ghanaian GDP per capita and the value of statistical life, and a series of explicit assumptions (Woods et al., 2016). This threshold is expressed as a range rather than a single estimate to reflect the uncertainty in the data used. Thus, to examine the potential cost-effectiveness of the HT-test in this study, the

threshold value of less than $104-$951/QALY was adopted, consistent with what decision makers currently use in Ghana.

1.4 Structure of the thesis

Following this introductory chapter, this thesis consists of six further chapters. Chapter 2 sets the context for this thesis by exploring and summarising the relevant literature related to the topic area. It presents the literature on test evaluation by highlighting the main types and applications of medical tests, test performance characteristics, the current focus of test evaluation and the paradigm shift from focussing on diagnostic accuracy to evidence on clinical effectiveness and cost-effectiveness. It further presents the literature on the methods, examples and challenges of economic evaluations with specific focus on the reasons for undertaking economic evaluations at the early phases of medical test development. Then the literature on testing in resource-limited settings is presented with specific focus on typhoid testing in Ghana and related issues.

Chapter 3 presents a systematic review focussed on examining the methodologies and tools that have been employed in early economic evaluation studies of medical tests. Specifically, it summarises how the problem of insufficient data for model parameterization in early modelling studies of medical tests has been managed, whether and how testing pathways have been modelled, and whether and how the uncertainty that accompanies early modelling studies has been acknowledged in sensitivity analysis.

Chapter 4 presents the collection of primary data. It shows how the qualitative methodology (framework analysis) was used to explore and delineate existing test-treat pathways for typhoid fever in Ghana. Specifically, it summarises the typhoid test-treat pathways in Ghana, existing

variations in care pathways and the reasons for the existing variations. It further presents the views of a range of health care professionals working in Ghana on the role of a hypothetical rapid diagnostic test for typhoid fever in the care pathway they work in.

Chapter 5 presents a systematic review focussed on exploring the literature on typhoid economic evaluations, the types of models that have been previously adopted for modelling typhoid interventions and capturing of data on model inputs that may be useful for a de novo model. Specifically, it summarises how interventions for typhoid fever have been evaluated.

Chapter 6 presents the methodology and results of an early cost-effectiveness study which examines the potential cost-effectiveness of the HT-test, with specific focus on the estimation of the maximum cost and the minimum test performance required for the test to be cost-effective. It explores the value of methodologies such as headroom analysis, the cost-effectiveness acceptability curve (CEAC) and value of information (VOI) analysis.

Chapter 7 brings together the findings of the thesis and discusses its relevance as a valuable source of information to technology developers on the market potential of developing the HT-test. It further discusses its relevance as a valuable source of information to policy makers to support their decision-making on the further development of a rapid test for typhoid diagnosis in Ghana, which could help allocate limited budgets more efficiently. It also discusses the findings of the study and its implications for policy changes that will greatly improve the delivery of care for typhoid fever patients in Ghana. Finally, it discusses the implications of the findings for the early economic modelling of tests.

Chapter 2: Literature review

This chapter sets the context for this thesis by exploring and summarising the relevant literature on test evaluation and economic evaluation of medical tests. First, medical tests are defined and their role in the health care setting is highlighted. Then, the literature on test evaluation is presented by highlighting the main types and applications of medical tests, test performance characteristics, the current focus of test evaluation, the paradigm shift from diagnostic accuracy to evidence on clinical effectiveness and cost-effectiveness. Furthermore, the literature on the methods, examples and challenges of economic evaluations at the early and late phases of product development are presented, with specific focus on the reasons for undertaking economic evaluations at the early phases of medical test development. Lastly, the literature on testing in resource-limited settings is presented with specific focus on typhoid testing in Ghana and related issues of test accuracy, cost, slow turnaround time and the lack of availability of current testing strategies.

2.1 Medical tests: purpose in the health care setting

A medical test is defined as "a characteristic, measurement, observation, or medical procedure made on a person that provides information that relates to some aspect of their health state'' (Deeks et al., 2014, p.5). As a concept, the application of tests can be seen primarily as an information gathering process in health care (Lee et al., 2010). The information gathered can be useful in a variety of ways including modifying the probability of the presence or absence of disease, assessing stages of disease progression, risk stratification, monitoring of chronic disease progression, monitoring a patient's condition following treatment and predicting future events (Knottnerus et al., 2002; MSAC, 2005). Tests are sometimes an integral part of treatment. For example, endoscopy is used not only to detect lesions but also to remove the detected lesions. As a result of the valuable role they continue to play in health care delivery

(Redekop, 2006), the number of tests has been increasing rapidly over the past decade and is likely to increase further (Koffijberg et al., 2013).

2.2 Main types of tests

The following describes the main types of tests or measures employed in health care settings to evaluate a patient's condition: medical history taking, physical examination, laboratory tests and pathology, physiological measurements and imaging tests.

Medical history taking involves a structured assessment carried out to elicit information that can provide a comprehensive evaluation of a patient's health status. It involves an assessment of the patient's current and previous health problems and medical treatment, factors which might affect the patient's health and their family's history. Taken together, the medical history, information from physical examination and the results of other tests are often used to produce a list of plausible diagnoses. The medical history of a patient is captured using questionnaires which may differ between settings.

Physical examination involves an examination of the body of a patient for signs that are diagnostic of a particular disease or complications of a particular disease. There are different tests that can be performed during a physical examination that depend on their rationale. For example, there is physical examination for abdominal tenderness and severe abdominal pain in typhoid fever as an indication for intestinal perforation.

Laboratory tests and pathology normally involve taking samples of blood, urine, stool or tissue and subjecting them to either chemical or mechanical analysis or by observation to look for abnormalities within the samples. In most cases, the side effects attributable to these tests are

those resulting from obtaining the samples (Stockham and Scott, 2013). Examples include urine or blood culture test, pap smear, blood glucose test and full blood count test.

Physiological measurements are made to assess how well the body functions. These may range from simple measurements of body temperature, blood pressure, weight and height to more complicated measures such as measuring heart functioning by taking an ECG (Webster, 2014).

Imaging tests are tests that produce images of parts of the body. The type of test requested will depend on the symptoms of the patient and the part of the body being examined. They include X-rays, Computed Tomography (CT) scans, Magnetic Resonance Imaging (MRI) scans, Ultrasound, positron emission tomography (PET) and endoscopic investigations.

2.3 Applications of tests

Tests can be categorised into their use for the purposes of screening, assessing predisposition, diagnosis, staging, prognosis, surveillance, clinical monitoring, assessing response and for stratification (Deeks et al., 2014). An appreciation of the various ways in which tests can be used is very important because, tests used in different ways may require different evaluative approaches. The applications of a test where management decisions are based on a test result at one point in time (i.e., the test is done and a management decision is made) fits the diagnostic evaluation paradigm. Whereas questions of longitudinal nature (i.e. tests that seek to predict a future event or management decisions are based on repeated test results or a change) do not fit the diagnostic evaluation paradigm. The test applications which fit the diagnostic evaluation paradigm are: diagnosis, screening, staging, and surveillance.

2.3.1 Diagnosis

Diagnosis is the process of establishing the presence or absence of a particular disease, condition or syndrome in a patient at the time of testing through the evaluation of the patient's history, review of laboratory data and examination of the individual (WHO, 2005). Tests are performed in patients who exhibit signs and symptoms indicative of the presence of a particular disease. The role of tests in diagnosis includes increasing or decreasing the probability of a particular condition being present, for staging or measurement of disease severity and for general examination for clues to the cause of symptoms (MSAC, 2005). An example of a diagnostic test is the Widal test used for the diagnosis of typhoid fever. To evaluate the potential value of a diagnostic test involves an assessment of its accuracy, how it influences diagnoses and therefore patient management, and ultimately how it improves outcomes relevant to the patient (see Fig 2-2). Summary statistics such as sensitivity, specificity, predictive values and likelihood ratios are all measures of test accuracy (Deeks et al., 2014). All of these terms are defined in later sections.

2.3.2 Screening

Screening involves testing asymptomatic patients to identify those people for whom further tests or treatments may be beneficial (Wilson and Jungner, 1968). Screening requires the use of accurate tests for early detection of conditions to allow for timely treatment and also the provision of information on risk to allow for efforts to modify the risk (Wilson and Jungner, 1968). Magnetic resonance imaging (MRI) is an example of a screening test.

2.3.3 Staging

Tests performed for staging purposes are implemented with the objective of establishing how advanced or severe a disease is. Such tests are performed in persons already known to have the disease leading to the identification of the optimal intervention for them. Staging tests further classify diagnosed patients into sub-diagnoses groups (Deeks et al., 2014). Staging answers the question, 'what stage of disease does the patient have'. An example of a staging test is the Oncotype DX test for breast cancer.

2.3.4 Surveillance

Tests performed for the purposes of surveillance are conducted to identify whether patients with established disease have progressed or regressed or had a recurrence. Detection of progression or recurrence may prompt changes in patient management. For example, patients who have received treatment for prostate cancer will have their prostate specific antigen (PSA) test measured to detect recurrence, and those with severe liver disease will have tests undertaken to assess progression of fibrosis.

2.4 Test evaluation: what is the current focus?

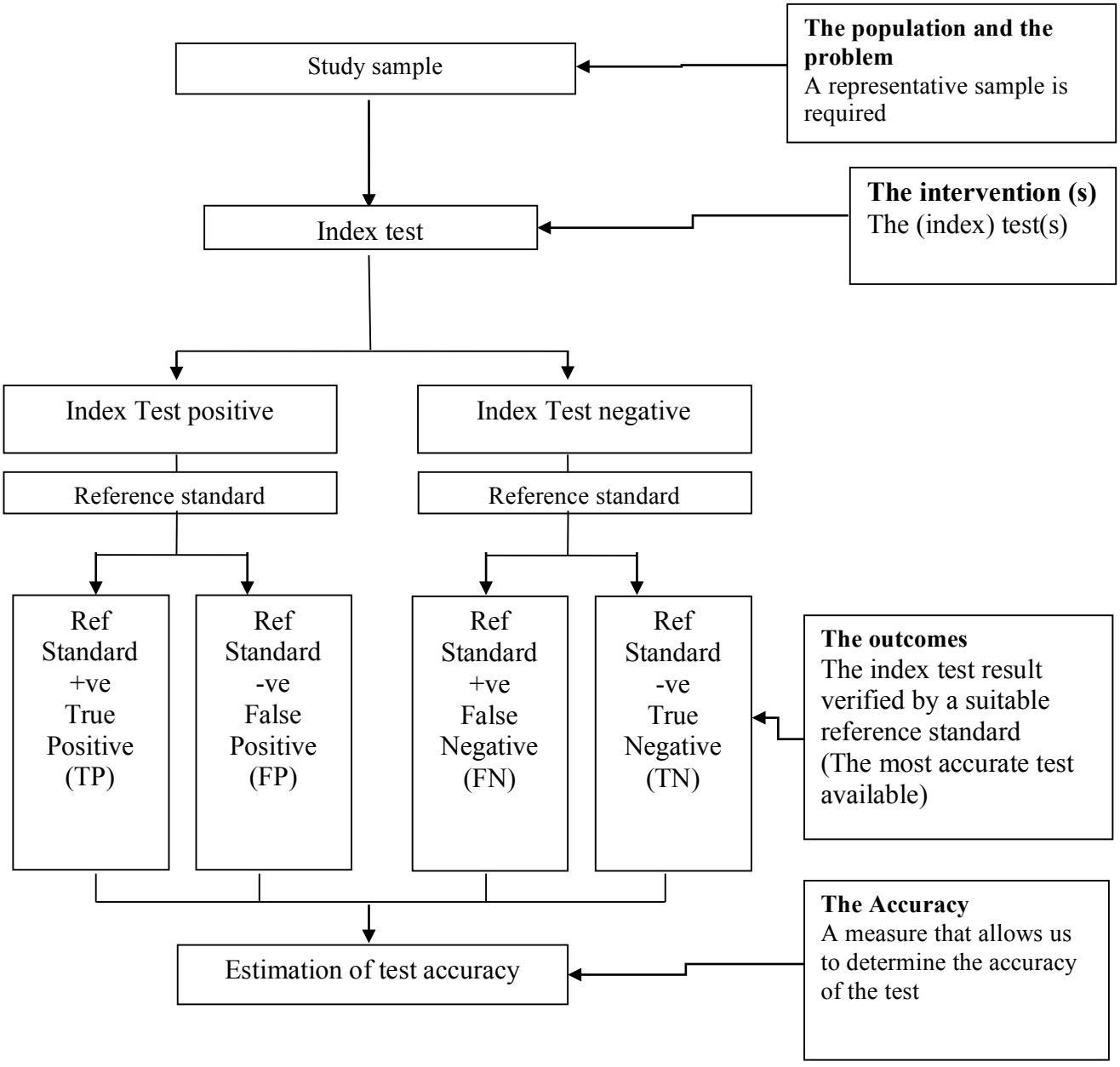
As previously described, the number of tests has been increasing rapidly over the past decade and is likely to increase further (Koffijberg et al., 2013), which is as a result of the valuable role they continue to play in health care delivery (Redekop, 2006). However, in an era of constrained government budgets and rising costs, this situation highlights the need for the proper evaluation of tests before their introduction into clinical practice. This will promote the adoption of innovative clinically and cost-effective tests.

Consequently, before a medical test is introduced in practice, payers, clinicians, policy makers and patients need information on the evaluative dimensions of the test. These evaluative dimensions include test performance (safety, analytical validity and reliability of the test), test accuracy (clinical validity; performance in 'suspected disease'), clinical effectiveness, cost, and cost-effectiveness. Although it is expected that all these evaluative dimensions are completely assessed before the introduction of a test into clinical practice, in reality, this is not so with the evaluation of test accuracy (not test effectiveness) predominating in test evaluation. The evaluation of tests typically stops after quantifying the accuracy of the test compared to the prevailing reference standard (Koffijberg et al., 2013).

2.4.1 What is test accuracy and how is it assessed?

Test accuracy is defined as the proportion of subjects that the test correctly identifies as being positive or negative for the presence of disease (Van den Bruel et al., 2007). The two most commonly reported measures of test accuracy are sensitivity and specificity (Gazelle et al., 2005). Summary performance characteristic measures include the diagnostic odds ratio and area under the receiver operator characteristic (ROC) curve (AUC) (Davenport et al., 2014). Other measures include positive and negative predictive values. In test accuracy studies, the results of the test under evaluation (index test), are compared with those of the reference standard (the most accurate currently available test) (see Fig 2-1 and Table 2-1) (Leeflang et al., 2008). Test accuracy is not fixed and may vary between patient subgroups and their spectrum of disease, the clinical setting, test interpreters and the result of prior testing (Leeflang et al., 2008). It is important to state that test accuracy is an explicit recognition that most tests make errors even if correctly performed.

Fig 2-1 Basic design to assess test accuracy of the target condition (condition we wish to detect)



The information obtained using the above study design is then used to estimate the metrics for expressing test accuracy using the 2x2 table.

2.4.2 Metrics for expressing test accuracy

Metrics for expressing test accuracy which as described in the previous section include, sensitivity, specificity and likelihood ratio, the diagnostic odds ratio, receiver operator characteristic (ROC) curve, positive and negative predictive values (Davenport et al., 2014). The information from a 2x2 table (Table 2-1) is used for calculating the measures of test accuracy.

Table 2-1 2X2 Table for calculating measures of diagnostic accuracy

	Patients with target disease	Patients without target disease
Positive index test result	True positive (TP) **a**	False Positive (FP) **b**
Negative index test result	False negative (FN) **c**	True negative (TN) **d**
TOTAL	**a+c**	**b+d**

2.4.2.1 Sensitivity

The sensitivity of a test is defined as the proportion of patients that have the target disease who test positive (Akobeng, A.K, 2006). It is calculated mathematically from the 2x2 table using the formula a/(a+c). If the estimated sensitivity is sufficiently high, the presence of disease is ruled out by a negative test result. A high sensitivity is particularly important if the consequence

of missing a disease has serious implications in terms of health outcomes for the patient and cost of treatment delays.

2.4.2.2 Specificity

The specificity of a test is defined as the proportion of patients without the target disease who will test negative (Parikh et al., 2007). This can be calculated from the 2x2 table as d/(b+d). If the estimated specificity is sufficiently high, a positive result rules in disease. A high specificity is particularly important if a false positive result can harm the patient (i.e., harms due to treatment and associated costs).

Advantage of sensitivity and specificity as an accuracy measure

It allows for the full implications of test accuracy to be examined in model-based economic evaluation studies. That is, it allows for each of the different possible test results to be modelled explicitly (i.e., TP, FP, FN, TN). Sensitivity and specificity are not directly affected by prevalence of the target condition and thus the results from one study may be applicable to different populations (although they may vary by disease severity which may in turn vary with disease prevalence).

2.4.2.3 Likelihood ratio (Positive Likelihood ratio (LR+) and Negative likelihood ratio (LR-))

A likelihood ratio estimates the probability of a test result (positive or negative) in patients with the target disease compared to those without the target disease (Thornbury et al., 1975). The positive likelihood ratio is defined mathematically as the probability of a "diseased" person testing positive divided by the probability of a "non-diseased" person testing positive. This is calculated as sensitivity/(1-specificity). The negative likelihood ratio is defined mathematically

as the probability of a "diseased" person testing negative divided by the probability of a "non-diseased" person testing negative. This is calculated as (1-sensitivity)/specificity.

A likelihood ratio greater than 1 increases the probability that the target condition is present whereas a likelihood ratio less than 1 decreases the probability that the target condition is present. A likelihood ratio of 1 indicates no change in the likelihood of disease and the test does not provide any useful diagnostic information.

2.4.2.4 Diagnostic odds ratio (DOR)

This is defined as "a measure of the odds of a positive test in those with the disease compared to those without the disease to show the association between a dichotomous test result and the diagnosis of the target disorder" (MSAC, 2005, p.49). The diagnostic odds ratio is calculated using the formula

(sensitivity/(1-specificity))/((1-specificity)/specificity) or (LR+/ LR-).

Estimates of diagnostic odds ratio ranges from 0 to infinity, with higher values indicative of better discriminatory test performance. A diagnostic odds ratio value of 1 indicates that the test does not discriminate between patients with or without the target condition.

2.4.2.5 Receiver-operator characteristic (ROC) curve

The ROC curve is a plot of test sensitivity (on the vertical axis) versus its false positive rate (1-specificity on the horizontal axis) at different thresholds (MSAC, 2005). The ROC curve demonstrates the trade-offs between the sensitivity and specificity of the test. The area under the ROC curve is a useful measure of the overall accuracy of the test. Values for the area under the curve (AUC) range from 0 to 1. An AUC value of 1 indicates that the test is perfect and the

sensitivity and specificity of the test is 100% at each cut-off point. An AUC value of 0.5 indicates that the test does not discriminate between the presence and absence of disease.

2.4.2.6 Positive Predictive Value (PPV) and Negative Predictive Value (NPV)

Positive predictive value represents the proportion of positively tested individuals who are diseased whereas negative predictive value represents the proportion of individuals with a negative test who do not have the disease. Although these are the most clinically informative measure of the accuracy of diagnostics, they are affected by prevalence of the target condition and thus cannot be readily transferred to different populations or pooled to produce a summary estimate (MSAC, 2005). PPV can be calculated from the 2x2 table using the formula a/ (a+b) and NPV is calculated using the formula d/(c+d).

2.4.3 Measures of analytical validity

The following describes the measures of analytical validity of a medical test: reproducibility, repeatability, precision and bias.

2.4.3.1 Reproducibility and repeatability

Reproducibility refers to the variation in measurements made on a subject under changing conditions (Bartlett and Frost, 2008). A valid statement of reproducibility requires specification of the conditions changed. The changing conditions may be due to different measurement methods or instruments being used, measurements being made by different observers, or measurements being made over a period of time (Bartlett and Frost, 2008).

Repeatability refers to the variation in repeat measurements made on the same subject under identical conditions (Bartlett and Frost, 2008). Thus, measurements are made by the same instrument or method, the same observer, and that the measurements are made over a short period of time over which the variable being measured can be considered to be constant. Variability in measurements made on the same subject in a repeatability study can then be ascribed only to errors due to the measurement process itself (Bartlett and Frost, 2008).

2.4.3.2 Precision and bias

A test method is said to be precise when repeated determinations (analyses) on the same sample give similar results (Schmidt, 2013). When a test method is precise, the amount of random variation is small. Precision is a function of random error and 'balances out' upon repetition (Schmidt, 2013).

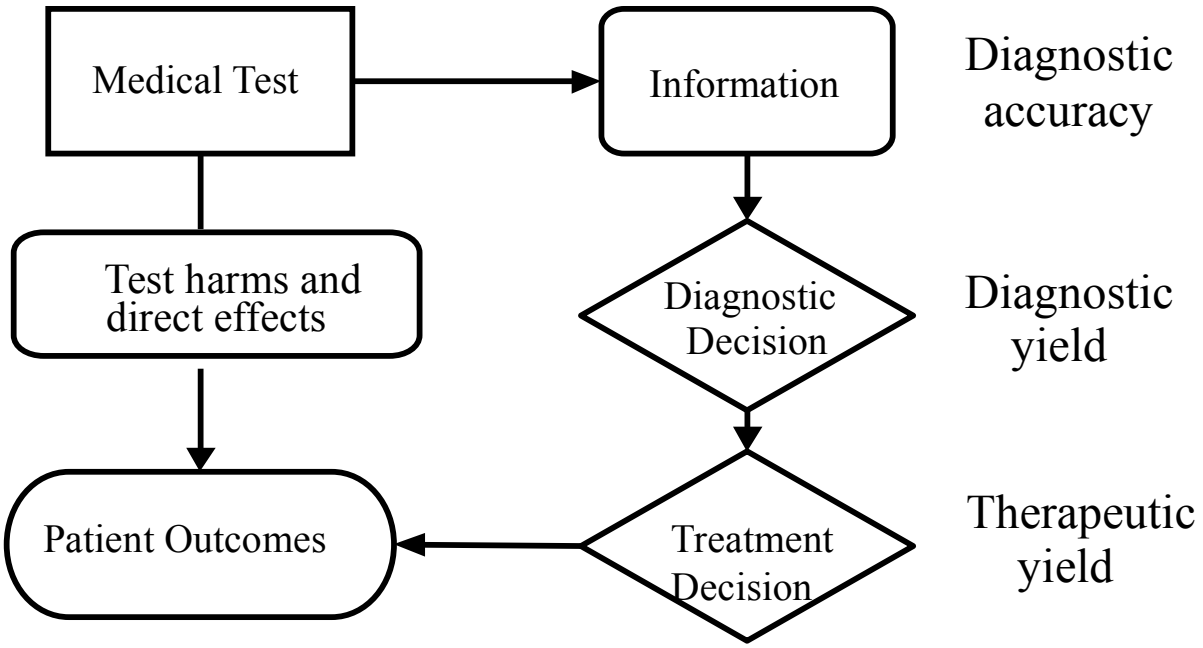
Bias is defined as a systematic difference in an observed measurement from the true value (Schmidt, 2013). In the context of test accuracy study, bias occurs when there is a systematic deviation from the true estimate of accuracy. If bias exists, a study would consistently overestimate or underestimate the true accuracy parameters if the study were repeated. Thus, bias is error that does not 'balance out' upon repetition (Schmidt, 2013).

2.5 Moving beyond Test accuracy: evaluating how tests impact on patient health on test-treatment pathways

Increasingly, decision makers, physicians and other users of medical tests request more than simple measures of test accuracy. They also want to know how testing leads to improved decision making and patient relevant outcomes, as well as the cost-effectiveness of testing to determine whether the test is good enough to use and represents value for money (Van den

Bruel et al., 2007). Moving beyond test accuracy, evaluation of tests should always be linked to the context of use (care pathway). This assertion supports the notion that tests are not neutral reporters of reality independent of context, rather, they are used in and are highly dependent on context (Matchar, 2012). Therefore, in order to establish the impact of a new test on relevant patient outcomes, it must be examined as part of a broader care pathway where the test is administered, and the test results are considered alongside other evidence to decide subsequent management decisions. It is only by considering this entire care pathway that the true value of a test can be determined, because changes to any stage of this pathway after the introduction of the test can impact on the outcomes experienced by patients (di Ruffano et al., 2012). The care pathway includes all the tests, decision making, and treatments associated with the use of the test both beneficial and harmful and their costs (NICE, 2011).

Fig 2-2 The test-treatment pathway



An in-depth understanding of the care pathway is also relevant for the following reasons. The introduction of a new test into clinical practice could serve as either a replacement test, a triage test or an add-on test (Bossuyt et al., 2006). Identification of its role in an existing care pathway is crucial in test evaluation, as it will determine which characteristics the new test should have. This is very important, as depending on the role of a test and its place in a pathway, maximising either sensitivity or specificity might be more important. Furthermore, the application of tests is associated with a series of outcomes (Lee et al., 2010); categorised into those resulting directly from the test, the testing process or both (AHRQ, 2012). Therefore, an in-depth understanding of the care pathway is key to the identification and selection of the relevant outcomes necessary for the evaluation of the test (especially in modelling). An a *priori* statement specifying the proposed care pathway between test results, treatment decisions, and patient outcomes is useful in the identification of the intended test population and comparator test strategy (distinct from reference standard) in the test evaluation process. Delineating the care pathway also helps clinicians, the majority of whom have difficulties in understanding test accuracy metrics (Whiting et al., 2015) to make a link between test accuracy and patient relevant outcomes. An understanding of the care pathway is very important in test evaluation and therefore, should be a rigorous undertaking involving all the necessary stakeholders such as clinicians, experts, and patients. Existing and emerging clinical guidelines may also provide a good source of information for this purpose.

2.5.1 Understanding the role of a test on the test-treat pathway

As previously described, a new test introduced into clinical practice can serve as either a replacement test, a triage test, or an add-on test (Bossuyt et al., 2006). Its role in the care pathway should be clearly defined and stated in order to enable the assessment of the test in terms of its use in the specific context.

2.5.1.1 Replacement test

A test intended to be a replacement test is expected to be at least one of: more accurate, less invasive, easier to interpret, quicker to yield results results compared to an existing test. For example, a replacement test could be of comparable accuracy, but its lower cost will make it a more cost-effective option.

Study design

To establish whether an index test can replace an existing test, it is important to compare the diagnostic accuracy of both tests. For a replacement test, a head-to-head comparison of the new and existing test is preferable, using a fully paired design (performing the new and existing test in each patient and comparing both to the reference standard) (Bossuyt et al., 2006). Indirect comparisons can also be conducted by randomly performing the new or existing test. Though indirect comparisons may be seen as a substitute to paired studies, they should be considered with great caution as the patient population, and reference standard should be the same in the studies that are being compared. Paired studies have the advantage of evaluating the two tests in the same patient group and it is also possible to use fewer patients (Bossuyt et al., 2006). In practice, head-to-head comparisons and RCTs are less common. Head-to-head comparisons may not be possible because of the burden to the patient of having to undergo more than one test. An example of where head-to-head comparisons are feasible would be comparison of 2 blood tests.

2.5.1.2 Triage test

In triage, the index test is used prior to the existing test or testing pathway, and only patients with a particular result on the triage test continue the testing pathway (Bossuyt et al., 2006). Triage tests are used to reduce the number of people who undergo the existing test, for example

because the existing test is invasive, expensive, difficult to access. In terms of accuracy, a sensitive triage test would be used to rule out the target condition prior to testing with the existing test (triage test negatives likely to be TN and do not receive further testing; triage test positives receive the existing test). Conversely, a specific triage test would be used to rule in the target condition and only negative triage test results would receive the existing test.

Study design

A test used for the purposes of triage does not aim to improve the accuracy of the current testing pathway. Rather it reduces the use of existing tests that are more invasive, cumbersome or expensive. Several designs can be used to compare the accuracy of the triage strategy. In a fully paired study design, all patients undergo the triage test, the existing test and the reference standard. As the primary concern is to find out whether disease will be missed and how efficient the triage test is, one option is to use a paired design and verify the results only of patients who tested negative on the triage test but positive on the existing test. This will identify patients in whom disease would be missed if the triage test is used as well as patients in whom the existing test can be avoided.

2.5.1.3 Add-on test

A test may be used as an add-on test as part of a clinical pathway if combination of 2 tests is more accurate than the accuracy of one test alone. For example, one test might have sensitivity greater than specificity and the add on test specificity greater than sensitivity for the target condition. Any increase in accuracy of add-on must be balanced against increased burden of testing for patient. Consequently, its application is reserved for specific patient subgroups in which the test is deemed a final resort. When a test is to be used as an add-on test, its necessary test characteristics depends on its goal. If the goal is to increase the number of patients to

receive treatment, then a high sensitivity is needed whereas if the goal is to increase the number of patients in which treatment is to be withheld, then a high specificity is required.

Study designs

More efficient methods with complete verification can be used to evaluate the effect of the add-on test diagnostic accuracy. For example, evaluating the accuracy of the existing test plus add-on test versus existing test alone. The difference in accuracy between the existing test testing strategy and the add-on test will depend exclusively on the patients who are positive on the add-on test. A study could therefore be limited to patients who were negative after the existing test but positive with the add-on test with verification by the reference standard. This design allows for the calculation of the extra true positives and false positives from using the add-on test.

2.6 The concept of economic evaluation

Economic evaluation is defined as "the comparative analysis of alternative courses of action in terms of their cost and consequences" (Drummond et al., 2005, p.9). Economic evaluations primarily seek to identify, measure, value and compare competing alternatives. Central to economic evaluation is the concept of opportunity cost (Morris et al., 2012). That is, by choosing to deploy resources to one use, the benefits of using those resources in their next best alternative are forgone. Therefore, it is important that health care resource allocation decisions are made such that there is value for money (WHO, 2000). Increasingly, economic evaluations are being used as a valuable tool in health care resource allocation both more explicitly and scientifically based, worldwide.

2.6.1 Types of economic evaluation

There are five types of economic evaluation (Morris et al., 2012). These are: cost-benefit analysis (CBA), cost-effectiveness analysis (CEA), cost-utility analysis (CUA), cost-consequence analysis (CCA) and cost-minimization analysis (CMA). These can be further categorized into partial economic evaluation methods (CCA and CMA) and full economic evaluation methods (CBA, CEA and CUA). In full economic evaluations, both cost and benefits are compared between two or more interventions. Partial economic evaluations consider costs and/or benefits, but either do not involve a comparison between competing interventions or do not relate costs to benefits. It is important to state that unlike full economic evaluations, partial economic evaluations do not provide information on efficiency (Morris et al., 2012). The main difference between the three types of full economic evaluation methods stems from the way benefits are measured and valued. In all three methods, cost is measured and valued in monetary terms. CEA measures benefits in terms of natural health outcomes (Jackson, 2012). For example, in terms of lives saved, complications avoided, symptom free days and ulcers avoided. CEAs are used to compare programmes with similar outcomes and the results are presented in cost per unit effect. For example, cost per successful typhoid treatment. Like any method, CEA has its own limitations. The use of a uni-dimensional outcome measure can be problematic because the total effects of a treatment may not be reflected in any one outcome measure and results cannot be compared when outcome measures are different. This may lead to inaccurate conclusions on the relative worth of the competing alternatives under investigation (WHO, 2000).

In CUA, benefits are measured in terms of quality-adjusted-life years (QALYs) gained or disability-adjusted life years (DALYs) averted (Kind et al., 2009; Murray et al., 2012). CUAs are used when health-related quality of life is the main outcome, interventions affect both

mortality and morbidity and the interventions affect a wide range of outcomes, but a single unit of comparison is required (Drummond et al., 2005). CUAs are presented in cost/QALY gained or cost/DALY averted. The limitation of this method of analysis as argued by some health economists is the implicit assumption that a "QALY is a QALY" regardless of who accrues it. However, as shown by Tsuchiya and Dolan (2009), society places different weights on QALYs gained across different population groups and thus "a QALY is not a QALY" regardless of who accrues it.

CBA attempts to measure benefits in monetary terms, thus, it is possible to determine whether the benefits of an intervention justify the cost. CBA is broader in scope than CEA and CUA. CBAs are presented in terms of net benefit (i.e., benefits minus cost). Its appeal over the other forms of analyses is that it provides a stronger basis for implementing an intervention or not because it can be seen whether the monetary value of benefits outweighs the cost (WHO, 2000). However, the limitation of CBA is that measuring and valuing benefits in monetary terms can be problematic.

It is worth stating that the three forms of full economic evaluation discussed above lend themselves to different decision rules. The choice to use a particular method of analysis should be made such that it is appropriate to the question(s)/objective(s) being addressed in the study. CEA and CUA seek to address technical efficiency questions. That is, seeking the best alternative that achieves a desired outcome at a minimised cost. Drummond et al (2005) puts it this way, "both CEA and CUA are techniques that relate to constrained maximisation; that is where a decision-maker is considering how best to allocate an existing budget" (Drummond et al., 2005, p.15). Whereas, CBA recommends choosing the alternative with the greatest net benefit.

2.6.2 Conducting economic evaluations

Economic evaluations are conducted in two main ways. They are either conducted alongside randomised controlled trials or they are conducted using decision analytic modelling. In trial based economic evaluation, cost and effectiveness data are collected simultaneously as the trial is being undertaken (Petrou and Gray, 2011). The advantages of doing this are that, it provides the opportunity for the researcher to collect data on resource use and cost pertaining to each patient in a study which may not be available at any other time and also, the cost of collecting economic data may be modest because it will be collected at the same time the trial is being conducted (Johnston et al., 1999). Despite these advantages, economic evaluations conducted alongside RCTs have limitations such as atypical patients/setting, truncated time horizons and limited comparators (Drummond et al., 2005). Decision analytic modelling is defined as "an analytic methodology that accounts for events over time and across populations, that is based on data drawn from primary and/or secondary sources, and whose purpose is to estimate the effects of an intervention on valued health consequences and costs" (Weinstein et al., 2003, pp. 10). Decision models can be used to structure the economic question and compare all relevant alternatives, extrapolate beyond observed data, link intermediate and final endpoints, generalize results to other settings and synthesise evidence and facilitate head-to-head comparisons where RCTs are not feasible (Drummond et al., 2005). The are two main types of models used in model-based economic evaluations namely: static and dynamic models. Static models used in health economics implicitly assume that the probability of disease exposure is constant over time and unaffected by interventions. For transmissible infectious diseases, this is not realistic and requires the use of dynamic models which allows for the explicit inclusion of transmission in the analysis (Lugner et al., 2010). There are two main types of static models namely: decision tree and Markov model. Decision trees are used for 'one-off' decisions and are particularly suited to acute care problems, once-only diseases and short term diagnostic or

screening decisions. Markov models represent disease processes which evolve over time, are suited to modelling the progression of chronic disease, can handle recurrence and can estimate long term costs and life years gained.

The concept of the economic evaluation is widely known for its application in the evaluation of pharmacological interventions and usually at the late phase of product development (Morris et al, 2012). Its application to medical tests is relatively new (including its application to tests at the early phases of development). Despite the increasing advocacy by many agencies that economic evaluations are best conducted alongside randomised controlled trials, this is not entirely the case in the cost-effectiveness analyses of medical tests. The reason is that test-treat randomised controlled trials capturing the downstream health outcomes arising from decisions made based on the test results are not feasible and are rare (Lord et al., 2009). Hence, decision analytic models are advocated as the most systematic and transparent method for the economic evaluation of medical tests to provide evidence on clinical and cost-effectiveness in the absence of test-treat RCTs (Trikalinos et al., 2009).

2.6.3 Decision analytic modelling of test-treat strategies: uses and issues

The following sections describe the uses and the issues associated with decision analytic modelling.

2.6.3.1 Uses of modelling in test-treat pathways

Evidence of test performance, treatment effectiveness and the prevalence of disease are reported in different studies across different populations and settings. Modelling provides a means of integrating this evidence from all these studies in a given setting to demonstrate

clinical effectiveness of a test. Furthermore, modelling allows for assumptions and uncertainties in key input parameters to be dealt with in sensitivity analyses to test the robustness of results when evaluating the clinical and cost-effectiveness of tests (Pauker and Kassirer, 1987). The application of all tests is associated with benefits, harms and cost and decision analytic modelling allows for such trade-offs to be assessed to aid in decision-making. Furthermore, modelling provides a framework for comparing multiple test-and-treat strategies and it also allow for exploring hypothetical conditions for diseases with no effective treatment by estimating under what conditions it would be worthwhile to employ such a test (Trikalinos et al., 2009). Decision models also provide the appropriate framework to extrapolate beyond trial evidence to a longer, more appropriate time horizon.

2.6.3.2 Modelling challenges faced in the evaluation of tests

Despite the advantages of using modelling in test-treatment pathways, it also has some challenges. There are challenges with insufficient/unreliable data, transferability of test performance across studies, choice of outcomes to model, and the parameterization of a model.

Issues with insufficient data

Excluding cost from this consideration, there are generally three main groups of important input parameters modelled in test-treat strategies (Trichinosis et al., 2009). These are prevalence of the disease in the setting of interest, test performance and the benefits and harms associated with the test-treat pathway. Problems arise when there is insufficient or unreliable data pertaining to any of these important input parameters as this can lead to an under or overestimation of incremental cost-effectiveness ratio resulting in the misallocation of scarce health care resources. Obtaining valid prevalence data especially for a particular setting or subpopulation is challenging but very important because prevalence significantly impacts on

the positive and negative predictive values of testing. Using unreliable or insufficient data on test performance can greatly affect estimates of sensitivity and specificity employed in modelling, which can consequently result in wrong estimates leading to a misallocation of scarce resources. Insufficient or unreliable data on treatment effectiveness would obviously also limit the validity of modelling results obtained on the clinical effectiveness of a test.

Issues with non-transferability of test performance data across studies

Estimates of sensitivity and specificity used in making calculations possible in models are often borrowed across studies. However, studies of test performance are not always performed in the setting of interest (Irwig et al., 2002), do not always evaluate the comparative accuracy of the test in its anticipated role (Bossuyt et al., 2006), and can evaluate the accuracy of different versions of the same test (Trikalinos et al., 2009). Hence transferring performance estimates across studies is problematic if no judgement calls are made on the appropriateness of these estimates before using them. Thus, because sensitivity and specificity are often considered not to be directly affected by prevalence of the target condition, these are frequently applied across settings with different prevalence estimates (Trikalinos et al., 2009). The challenge here is that differences in the inclusion criteria of different studies can result in spectrum effects (differences in the sensitivity or specificity of a diagnostic test according to the patient characteristics or disease features) hence highlighting the need for careful consideration when transferring test performance estimates across studies. Depending on the intended role of a medical test on the care pathway, the accuracy of the test is determined in that specific context. Therefore, one has to be careful when transferring performance estimates across studies that evaluate the test in different roles because performance estimates may differ depending on the intended role of the test. Ignoring this and generalizing estimates of test performance will lead to estimates that are not a true representation of the real value of the test under evaluation.

Furthermore, test performance can differ for different versions of the same test, thus when transferring test performance modellers should consider whether it is for the same version of a test or not, as using different estimates can provide misleading conclusions and have negative consequences on policy decisions.

Issues with the choice of relevant outcomes to be modelled

A key issue for trials and decision models is the selection of outcomes that need to be measured or modelled to evaluate how tests are affecting patients (di Ruffano et al., 2012). Key outcome categories include clinical management outcomes, direct health effects, emotional, social, cognitive and behavioural responses to testing (Bossuyt and McCaffery, 2009) and intermediate outcomes (diagnostic test accuracy statistics). In order to comprehensively evaluate the value of diagnostics, the full range of outcomes associated with their use should be examined and selected based on their effect on the care pathway. If the full range of outcomes is not explored from the outset, the likelihood of excluding important outcomes is high. Consequently, recommendations made about tests may not be appropriate in those circumstances. Some useful suggestions have been made for choosing relevant outcomes and a two-stage approach has been proposed (AHRQ, 2012). The first stage is to list outcomes methodically and the second stage is to solicit input from stakeholders. At the first stage, outcomes can be separated into those attributable to the testing process and those attributable to knowledge of the test results. Outcomes attributable to the testing process are the direct effects of testing whereas outcomes attributable to the test results are numerous and include the patient's response to the test results and how patients and clinicians act upon the results. The relative importance of these outcomes varies depending on the intended purpose of the test (screening, diagnosis or prognosis) and researchers should be aware of this in order to aid in their selection of outcomes. For example, the adverse emotional effects of testing will be more

significant for screening tests than for diagnostic tests due to the high proportion of false positive test results associated with many screening tests.

Issues with incorporating the value of testing unrelated to treatment effects

An important yet most overlooked subject in the assessment of the value of diagnostics is the value of testing unrelated to treatment effects. Many studies evaluating the value of diagnostics using the conventional cost-effectiveness analysis framework have typically valued diagnostics at the extent to which they improve medical decision making and outcomes directly related to treatment effects (Lin et al., 2013). This is worrisome because patients' perceptions of testing, their experience of the testing process, and their understanding of test results can all affect downstream health. Asch et al. (1990) observed that clinical impact alone is an insufficient measure of value for diagnostics because diagnostics also have the potential to affect patients' psychological wellbeing whether or not they affect treatment. For example, a test for dementia may have relatively little impact on treatment or outcomes but may have substantial effects on a patient's psychological wellbeing (Lee et al., 2010). Furthermore, a genetic test that can determine with certainty whether one would eventually contact Huntington's disease may be considered to have little value from the medical perspective because it does not affect treatment. However, it could be highly valued by a patient concerned with childbearing if she is at high risk for the disease based on family history. Patients may value or have preference for test information unrelated to treatment effects even in the absence of immediate treatment (Neumann et al., 2010). A study by Schwartz et al. (2004) in the USA found that 87% of adults favoured cancer screening. Neumann et al (2010) also found that most respondents were willing to take tests predictive of future cancer, Alzheimer's disease, and arthritis. In both instances, tests were favoured even though the respondents were made aware from the outset that there were no immediate treatments available after testing. Their preference stems from

several reasons. Test information may reduce uncertainty and provide reassurance, assist in life-planning decisions and enable them to live differently (Lin et al., 2013). Clearly, patients may value test information even when it does not lead to changes in treatment.

Incorporating the value of testing unrelated to treatment effects into economic evaluations could significantly impact on estimated benefits and their allocation. These effects of testing could accumulate to outweigh any clinical benefits of testing or may bring additional benefits adding to the clinical effects of testing (Bossuyt and McCaffery, 2009). Despite the potential impact of the value of testing unrelated to treatment effects, it has typically not been included in previous economic evaluations. This may be because collecting such information is seen as an expensive and time-consuming endeavour, its value may be considered too small to have a meaningful impact on the assessment, researchers are simply unaware of these effects, and there are methodological complexities associated with their measurements (Lee et al., 2010). Failure to recognise the importance of "patient preference" for test information unrelated to treatment effects runs a risk of under or overestimating the true value of tests, leading to a misallocation of scarce health resources. As the number of new diagnostics proliferates, this issue will only become widespread.

Issues with model parameterization

Although models are considered a simplified representation of what occurs in the real world (Trikalinos et al., 2009), they should be comprehensive enough to capture all the necessary components of the simulated situation. Parameterization of a model should always be done based on an in-depth understanding of the real-world situation (care pathway). The challenge here is that, modellers might not be intimately familiar with the intricacies of the topic area and getting clinicians (or experts) in the field to verbalise the diagnostic decision-making process

and the wider care pathway more generally can be difficult. The result is that modellers might miss important information that should be included in the model structure and parameterization and hence generate results that are not very representative of what happens in real life. Such situations can lead to a misallocation of scarce health care resources highlighting the need for further work to be conducted in this area.

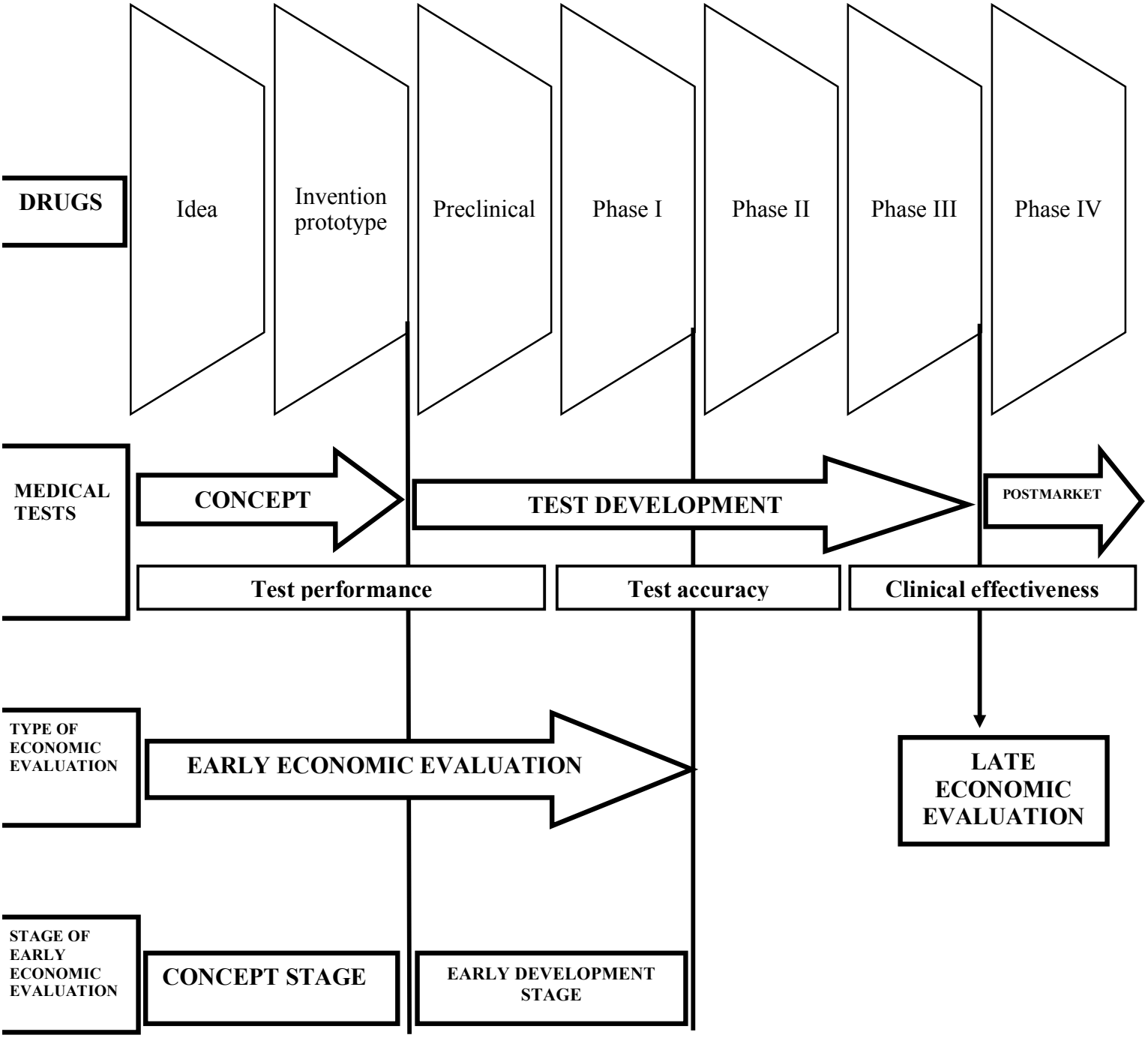
2.7 Economic evaluation of medical tests at the early phases of development

Figure 2-3 shows the various phases on the medical technology innovation pathway and the stages at which economic analyses are conducted (i.e., early and late phases). The early phase of development is categorized into two stages: the concept stage and, the early development stage. The concept stage is primarily the discovery and ideation phase. At this stage, the product is still hypothetical and in the case of medical tests, there is no available data on test parameters (e.g. accuracy, cost, etc.). The early development phase occurs between the equivalent of phase I clinical trials of drugs and the end of the concept stage. During this stage, certain test properties are assessed, and some form of experimental data may be available.

Typically, economic evaluations are conducted at the late stage of product development (Pietzsch and Paté-Cornell, 2008). For example, they may be done after Phase III of drug trials (Figure 2-3). The rationale is that this is the point at which there would be enough data on the product for its proper evaluation (Drummond et al., 2005). In the case of medical tests, it is expected that at this point test performance, test accuracy and clinical effectiveness would have been assessed. Test performance can be assessed at the preclinical stages, test accuracy in phases I and II, and clinical effectiveness in phases III and IV (Figure 2-3). Coverage and reimbursement decisions are therefore typically made on medical tests at the late phase of

development when a substantial amount of resource has already been committed to their research and development. This implies that any negative coverage and reimbursement decisions made on a test at this point would lead to no returns on investment and a loss being incurred (loss to manufacturers; opportunity cost for research; and inefficient use of resources allocated to Health Technology Assessment). Typically, advances in medical testing technology occur more frequently than for drugs (NICE, 2011). Therefore, leaving an economic evaluation to late stages of development may make any new findings redundant. Thus, there is now a growing interest in the economic analyses of medical tests at their early phases of development by investors, innovators, and policy makers, to identify their potential economic value and likely impact (Buisman et al., 2016; Tu et al., 2014).

Figure 2-3 The economic evaluation on the medical technology innovation pathway



2.7.1 Early economic evaluation: definition and characteristics

The early economic evaluation of a medical technology is defined as an iterative economic evaluation process to assess its economic value and likely impact (Buisman et al., 2016) at a time where it can still be considered experimental or emerging (Hartz and John, 2009). In the case of medical tests, this is usually from the concept stage up to stage I clinical trials (Markiewicz et al., 2014). Early economic evaluation provides useful information that can inform investment and design decisions under conditions of high uncertainties before the clinical performance of the test is actually established (Pietzsch and Paté-Cornell, 2008).

A key characteristic of early economic evaluation is the use of limited data, which is associated with increased uncertainty (Tappenden et al., 2004; Tuffaha et al., 2014). Value of information (VOI) analysis is recommended when there is considerable amount of uncertainty associated with data (as will be the case for early economic evaluations) (Briggs et al., 2006). VOI analysis is recommended because it is an important tool that decision makers can use to decide whether the available evidence used in an analysis is sufficient or not to fund the further development of a test. It can also be used to identify specific areas where further information is needed to decide on further development decisions (Chen and Willan, 2014). This will ensure the efficient use of limited resources and potentially assist to avert investing in a test that is not economically viable. Evaluating the potential value of any new technology at its early stage of development would require the use of models because of the lack of data describing its effectiveness during its early stages of development. Models are associated with the ubiquitous existence of uncertainty in their structure and model parameter estimates. This situation highlights the need to include sensitivity analysis in early economic evaluation to test the robustness of parameter estimates, determine the range of parameters which have the greatest

impact on cost-effectiveness, and determine how sensitive the results are to changes in model structure (Morris et al., 2012).

Another key issue in early economic evaluation is the "maximum cost" at which the new testing strategy is still cost-effective at a given willingness-to-pay (WTP) threshold. This "maximum cost" is known as the headroom (McAteer et al., 2007). In the headroom paradigm, the price at which the net monetary benefit of using the new technology equals the net monetary benefit of using the comparator technology is the headroom price. The estimated headroom provides valuable information that would guide decision making on the feasibility of further product development (in this case further test development). Thus, its consideration at the early stages of development especially the concept stage will promote the efficient use of limited resources. The decision rule in the headroom paradigm is that, if the estimated headroom is too low and realistically it is not possible to develop the new technology below this price, resources should not be committed to its further development. The WTP threshold at which the headroom of the new technology is determined is informed by the country within which the product will be adopted for use. For example, a new technology intended for use in the UK will be assessed at the WTP threshold of £20-30,000 (McCabe et al., 2008).

.

2.7.2 Early economic evaluation of medical tests: importance to decision makers and innovators

Economic evaluations conducted at the early phases of medical test development have potentially profound advantages for both decision makers and innovators. For decision makers, results of an early economic analysis could help allocate limited budgets more efficiently by identifying which tests to fund for further development, and which tests to reimburse on condition of further data collection (known as "coverage with Evidence Development" as

practiced in the USA and UK) (Hartz and John, 2009). Furthermore, early economic analyses studies speed up decision making regarding test adoption in the late phase of development, and support the management of test diffusion through "horizon scanning": the early identification of new economically viable medical devices (Hartz and John, 2009). For innovators, early economic evaluation helps to estimate the maximum cost of a new test and the minimum test performance required for the test to be cost-effective at a given willingness-to-pay threshold. In addition, it provides valuable information for making decisions on further test development (Grabowski, 1997), designing and managing reimbursement strategies, setting realistic performance-price goals and guiding decisions on resource investment in the test development process (Hartz and John, 2008). Clearly, early economic evaluation of medical tests has profound advantages for both decision makers and innovators and has become increasingly important as there is the demand to demonstrate value for money (Buisman et al., 2016).

2.7.3 The similarities and differences between early and late health technology assessment (HTA)

The similarities and differences between economic evaluations conducted at the early and late phases of a medical test development are summarised in Table 2-2 below. It provides information on the aims, the level of evidence used and how early and late HTA assessments are used to support decision making.

Table 2-2 Similarities and differences between early and late HTA

	Traditional HTA	**Early HTA**
Aim	To evaluate the safety, effectiveness and cost-effectiveness of a new medical technology	To evaluate the potential safety, effectiveness and cost-effectiveness of a new medical technology
Decision support	To support payers, regulators and patients on payment, clearance and test usage decisions respectively	To support innovators and investors on decisions of further development, designing and managing reimbursement strategies
Available evidence	Clinical studies of new technology	Earlier version of the medical technology, clinical experience and expert opinion
Influence on technology performance	Normally does not influence the clinical performance of the technology although it can occasionally have limited influence	It can potentially influence greatly the performance of the medical device in the future

2.8 Medical testing in resource-limited countries

Despite the potentially valued role medical tests continue to play in health care delivery worldwide, many of these tests are centralised, need highly trained staff and specialised facilities, are generally expensive and require skilled technicians for their regular maintenance (Drain et al., 2013). This has resulted in standard laboratory tests being cost prohibitive and inaccessible in resource-limited countries leading to avoidable morbidity and mortality from ailments that would have otherwise be easily treated (Peeling et al., 2010). In recognition of this disparity, the World Health Organisation (WHO) and other agencies have called for new diagnostic methods that can function in resource-limited settings (Urdea et al., 2006). This is because it has been estimated that more than 1.2 million deaths each year could be prevented in resource-limited countries through the deployment of rapid diagnostic laboratory independent tests for only four infections; syphilis, tuberculosis, bacterial pneumonia and malaria (RAND Corporation, 2007). For these tests to be useful in these settings, they have been accurate and available with a quick turnaround time that has enabled prompt implementation of appropriate effective therapy. There are studies that have demonstrated "value for money" of rapid diagnostic testing in resource-limited settings (Ansah et al., 2013; Hansen et al., 2015; Mallma et al., 2016). It is therefore not surprising that there has been an increase in the popularity and potential application of rapid diagnostic laboratory independent testing in resource-limited countries (Schito et al., 2012). Despite this increase, priority setting regarding test evaluation and development is not effective and does not reflect priorities in individual countries (Odame, 2013; Schroeder et al., 2015). Consequently, some disease areas that pose serious public health threats and increasing cost to health care systems are neglected. One such neglected disease is typhoid fever (which is the case in the resource-limited country Ghana) (Wain et al., 2015).

2.8.1 Typhoid testing in Ghana and related issues

Typhoid fever is an acute infection caused by the bacteria *Salmonella typhi* and *paratyphi* (incubation period 3-60 days) (WHO, 2003). Humans serve as the only natural host and reservoir for the causative organisms. Transmission is via the consumption of food or water contaminated with the pathogens and not direct person-to-person (Akullian et al., 2015). Typhoid fever remains an important cause of disease worldwide, but particularly in resource-limited countries (Keddy et al., 2011), because of inadequate diagnostic laboratory capacity to differentiate typhoid fever from other febrile conditions (Wain et al., 2015), and poor hygiene practices (Tilahun et al, 2017). In Ghana, of a population of about 29 million, the population with adequate sanitation facility access is, total: 14.96%, urban: 20.2%, rural: 8.6% and inadequate sanitation facility access is, total: 85.1%, urban: 79.8%, rural: 91.4% (CIA, 2015). Thus, the country is high risk for typhoid fever. Once ingested, *Salmonella typhi* multiplies and spread through the blood stream to all organs. Typhoid fever can be a serious illness with symptoms usually developing between 1-3 weeks after ingestion and may be mild or severe (WHO, 2003). The symptoms of typhoid fever include high fever, fatigue, weakness, stomach pains, constipation or diarrhoea and loss of appetite. In some cases, patients have rose-coloured spots on the chest, enlarged spleen and liver (GNDP, 2014). If left untreated, intestinal perforation with peritonitis may occur, resulting in a sudden drop in blood pressure leading to shock and death. Acute psychosis and severe intravascular haemolysis leading to acute renal failure may also occur. Typhoid fever is associated with an estimated case fatality rate of approximately 1% with prompt appropriate antimicrobial therapy but approaches 30-40% after intestinal perforation, which occurs in 1-3% of hospitalized patients (Neil et al., 2012).

2.8.2 What are the issues of concern with typhoid testing?

In Ghana, there are issues of concern with the diagnosis of typhoid fever such as delays in diagnosis, concerns about the accuracy of current tests, and lack of availability (see Chapter 1, p.1-2). All of which highlight the need for the development of a rapid, accurate and easily accessible diagnostic test in Ghana.

The availability of such a test will improve prompt diagnosis of typhoid fever, promote rational and judicious antibiotic prescribing practices by health professionals in Ghana (which is a major cause of drug resistance), and provide accurate surveillance data on disease burden that will aid in monitoring and preventing outbreaks. Furthermore, the availability of such a test can save the Ghanaian economy huge sums of money reimbursed for the unnecessary treatment of typhoid fever in patients without the disease, and treatment of patients with complications of typhoid fever because of delayed or overlooked diagnosis. Furthermore, monies that would be spent during outbreaks or on the research and development of effective treatments for resistant strains can also be saved and used in other disease areas to improve the health and well-being of Ghanaians. The added benefit of testing in this area cannot be relegated to the background because of the serious implications it has both on the health of Ghanaians and on the economy of Ghana. This has resulted in the urgent need to develop a cost-effective rapid diagnostic test for typhoid fever in Ghana, thus, the relevance of this study.

2.9 Conclusion

Having discussed the relevant issues, concepts and principles that relate to the topic area, highlighting the need for the development of a rapid diagnostic test kit for typhoid fever in Ghana, the next chapter (Chapter 3) is focused on examining how early economic evaluations

of medical tests have been conducted to date. This helps to gain greater insight into the methodologies and tools that have been employed in early evaluations of medical tests and thereby ensure that the most robust methodologies are adopted in the evaluation of the hypothetical test described in Chapter 6.

Chapter 3: Economic evaluation of medical tests at the early phases of development: a systematic review of empirical studies

This chapter focussed on examining the methodologies and tools that have been employed in early economic evaluation studies of medical tests to answer the question, how have economic evaluations conducted during the early phases of test development been done to date, and can any lessons be learnt from them? This chapter seeks to provide a greater understanding of the methodologies and tools that have been employed in early evaluations of medical tests and to inform the early economic analysis described in Chapter 6.

3.1 Background

There is an increasing demand to demonstrate value for money in health care systems worldwide and early economic evaluations are becoming increasingly important in this regard. As discussed in Chapter 2, early economic evaluations of medical tests have potentially profound advantages for both decision makers and innovators and their relevance in decision-making cannot be ignored. As a result of their role in promoting efficiency, there is now increasing interest in the economic analysis of medical tests at the early phases of development. However, despite this increasing interest there is little specific guidance on its implementation (Buisman et al., 2016). Early economic evaluations are more iterative in nature, are conducted at a time when there is a lack of data, and thus, are associated with increased uncertainty compared to late economic evaluations. This is further compounded by the increased uncertainties associated with mapping clinical pathways in test evaluation (see Chapter 1, p.3). These differences suggest that the methods used in economic evaluations conducted at the late

phase of medical test development are not applicable and need modification for use at the early phases.

3.2 Study aim and objectives

The aim of this systematic review was to examine the methodologies and tools that have been used in previous early economic evaluation studies of medical tests. Specifically:

1. To gain a greater understanding of how the problem of insufficient data for model parameterization has been managed.
2. To understand whether and how testing pathways have been modelled.
3. To explore whether sensitivity analysis has acknowledged the uncertainty that accompanies early modelling and the stage at which it is conducted.

3.3 Methods

For this systematic review, the following databases were searched for studies published from inception to July 2016. No language restrictions were applied.

1. Cochrane Library (CENTRAL)
2. Medline
3. EconLit
4. Centre for Reviews and Dissemination [Database of Abstracts of Reviews of Effects (DARE), Heath Technology Assessment (HTA), and NHS Economic Evaluation Database (NHS EED)]
5. Health Management Information Consortium (HMIC)
6. Excerpta Medica dataBASE (EMBASE)

Internet searches were also conducted (e.g., Google scholar and the websites of relevant organisations related to innovations in health care such as EuroScan). This was done to identify

any grey literature and references missed by the database search. The nature of early economic evaluations means that they are less likely to be published. Therefore, seven medical technology companies were contacted to identify evaluations that have not been published to complement the literature review. A reference search was undertaken by scanning the bibliographies of the relevant articles finally included in the review for additional relevant articles. The reference management software Endnote was used to manage the list of articles.

3.3.1 Search terms

The search strategy was customized for each database. A combination of MESH terms and index terms were used as shown in Table 3-1 below. To ensure that all relevant studies on the topic area were retrieved, the search strategy was piloted after it was developed in consultation with an information specialist. The complete search strategy including how all the different question elements were combined is shown in Appendix 1.

Table 3-1 Question elements and terms used for searching

Question element	**Terms used for searching**
Time (early) aspect of the search	new or novel or emerging or innovati* or early earl* or mid* or develop* or formative* or determinative* or design* or concept* or investigation*) adj2 (phase* or stage* or process* or "life cycle" or lifecycle) nascent or original or "ground breaking" or groundbreaking or promising or "cutting edge" or "cutting-edge" or seminal or "recently introduced") early adj3 (valuation* or evaluation* or assessment* or model*) (early-stage adj3 (valuation* or evaluation* or assessment* or model*)). early adj2 (economic or estimate*)
Economic evaluation aspect of the search	models, economic/ or models, econometric/ (headroom or "head room"). markov chain/, Decision Trees/, decision support techniques/, (discrete event* adj8 model*), (decision* adj5 model*), (model* adj5 markov*), (econom* or cost or costs) adj model*), decision making/, cost benefit analysis/, technology assessment, biomedical/
Medical test aspect of the search	test*, exp "diagnostic test"/, exp screening/, exp prognosis/, exp "predisposition test"/

3.3.2 Inclusion criteria

A primary study was deemed potentially relevant to be included in the review if:

1. It was an economic evaluation (CEA, CUA, CBA, CCA or CMA) conducted at the early phases of development (early phase as defined in Chapter 2, p. 38); and
2. The evaluated technology was a medical test or series of tests used together (at least one test needed to be present in at least one arm of the analysis); and
3. A decision model was used in the analysis.

3.3.3 Exclusion criteria

Studies were excluded from the review if they were:

1. Trial protocols or commentaries; or
2. Letters or editorials.

3.3.4 Selection of articles for the review

After the removal of duplicates, a two-stage approach to article selection was used. The first stage involved screening of abstracts and titles to identify potentially relevant articles whilst the second stage involved screening of the full text articles. The screening process was undertaken by two reviewers (S.F and P.B) independently and any disagreements regarding eligibility were resolved using the opinion of a third reviewer (C.D) where necessary. The screening process during both stages was conducted against the inclusion and exclusion criteria. All studies identified after the second stage of article selection were included for data extraction.

3.3.5 Data extraction

A data extraction form was developed in Excel and piloted. Then data extraction was conducted for each study included in this review to answer the following questions:

1. How was the problem of insufficient data for model parameterisation managed?
   - What sources (type) of data were used?
   - Was the source (type) of data used influenced by the stage of evaluation?
2. Modelling of testing pathways
   - What was the type of model used in modelling disease progression and the testing pathways (e.g. decision tree, Markov)?
   - Were test accuracy and all possible test results considered (i.e. true positives (TP), true negatives (TN), false negatives (FN), and false positives (FP))?
   - Were all the issues on the test-treat pathway that may be important to patients included in the analysis (e.g., personal costs incurred when accessing testing, effect of testing pathway on quality of life, value of knowing)?
   - Did the analysis include headroom analysis, and at what stage of the analysis was it included?
3. Uncertainty
   - Was the uncertainty that accompanies early modelling acknowledge in sensitivity analysis?
   - Was value of information (VOI) analysis conducted?

3.3.6 Quality assessment

The methodological quality of included studies was assessed using a 10-point checklist for economic evaluations (Drummond and Jefferson, 1996). Table 3-2 below shows the quality

assessment tool employed in the study. A score was assigned based on how well criteria were met; scores of 1, 0.5 and 0 were assigned to "yes", "cannot tell" and "no" respectively. Thus, each study scored from worst (0) to best (10) (Gonzalez-Perez, 2002). However, because the focus of this systematic review was to explore the methodologies used and not to comment on the validity of results and conclusions drawn from these studies, none of the studies were rejected based on the quality assessment undertaken. It is worth stating that some level of caution is needed when scoring studies because items are not always equivalent and different methods are likely to produce different scores and in some instances, scoring may even bias quality assessment. However, in this methodological piece of work spanning a range of clinical areas where the purpose of quality assessment was concerned with the presence or absence of a methodological approach (i.e., descriptive rather than decisions about inclusion or exclusion or scoring), the potential limitations of quality scores were not considered applicable.

Table 3-2 10-point checklist for assessing economic evaluations

1. Was a well-defined question posed in answerable form?
    - 1.1. Did the study examine both costs and effects of the service(s) or programme(s)?
    - 1.2. Did the study involve a comparison of alternatives?
    - 1.3. Was a viewpoint for the analysis stated and was the study placed in any decision-making context?
2. Was a comprehensive description of competing alternatives given?
    - 2.1. Were there any important alternatives omitted?
    - 2.2. Was (should) a do-nothing alternative be considered?
3. Was the effectiveness of the programmes or services established?
    - 3.1. Was this done through a randomised, controlled clinical trial? If so, did the trial protocol reflect what would happen in regular practice?
    - 3.2. Was effectiveness established through an overview of clinical studies?

3.3. Were observational data or assumptions used to establish effectiveness? If so, what are the potential biases in results?

4. Were all the important and relevant costs and consequences for each alternative identified?

4.1. Was the range wide enough for the research question at hand?

4.2. Did it cover all relevant viewpoints? (Possible viewpoints include the community or social viewpoint, and those of patients and third-party payers. Other viewpoints may also be relevant depending upon the particular analysis.)

4.3. Were the capital costs, as well as operating costs, included?

5. Were costs and consequences measured accurately in appropriate physical units?

5.1. Were any of the identified items omitted from measurement? If so, does this mean that they carried no weight in the subsequent analysis?

5.2. Were there any special circumstances (e.g., joint use of resources) that made measurement difficult? Were these circumstances handled appropriately?

6. Were costs and consequences valued credibly?

6.1. Were the sources of all values clearly identified? (Possible sources include market values, patient or client preferences and views, policy-makers' views and health professionals' judgements)

6.2. Were market values employed for changes involving resources gained or depleted?

6.3. Where market values were absent (e.g. volunteer labour), or market values did not reflect actual values (such as clinic space donated at a reduced rate), were adjustments made to approximate market values?

6.4. Was the valuation of consequences appropriate for the question posed (i.e. has the appropriate type or types of analysis – cost-effectiveness, cost-benefit, cost-utility – been selected)?

7. Were costs and consequences adjusted for differential timing?

7.1. Were costs and consequences that occur in the future 'discounted' to their present values?

7.2. Was there any justification given for the discount rate used?

8. Was an incremental analysis of costs and consequences of alternatives performed?

8.1. Were the additional (incremental) costs generated by one alternative over another compared to the additional effects, benefits, or utilities generated?

9. Was allowance made for uncertainty in the estimates of costs and consequences

9.1. If data on costs and consequences were stochastic (randomly determined sequence of observations), were appropriate statistical analyses performed?

9.2. If a sensitivity analysis was employed, was justification provided for the range of values (or for key study parameters)?

9.3. Were the study results sensitive to changes in the values (within the assumed range for sensitivity analysis, or within the confidence interval around the ratio of costs to consequences)?

10. Did the presentation and discussion of the study results include all issues of concern to users?

10.1. Were the conclusions of the analysis based on some overall index or ratio of costs to consequences (e.g. cost-effectiveness ratio)? If so, was the index interpreted intelligently or in a mechanistic fashion?

10.2. Were the results compared with those of others who have investigated the same question? If so, were allowances made for potential differences in study methodology?

10.3. Did the study discuss the generalisability of the results to other settings and patient/client groups?

10.4. Did the study allude to, or take account of, other important factors in the choice or decision under consideration (e.g. distribution of costs and consequences, or relevant ethical issues)?

10.5. Did the study discuss issues of implementation, such as the feasibility of adopting the 'preferred' programme given existing financial or other constraints, and whether any freed resources could be redeployed to other worthwhile programmes?

Source: Drummond and Jefferson (1996)

3.4 Results

After the removal of duplicates, 4,494 unique articles were identified for the first stage of screening (title and abstract screening). Eighty-eight (88) titles and abstracts were identified as potentially eligible for inclusion. After the second screening stage (full text screening), and unsuccessful attempts to contact medical technology companies, manufacturers of the technologies identified from the EuroScan website and obtain full text conference abstracts, five studies were included for narrative synthesis. The PRISMA flow diagram (Fig 3-1) below illustrates the results of the paper selection process with reasons for exclusion noted.

Fig 3-1 The PRISMA flow diagram summarising the search strategy and the paper selection process

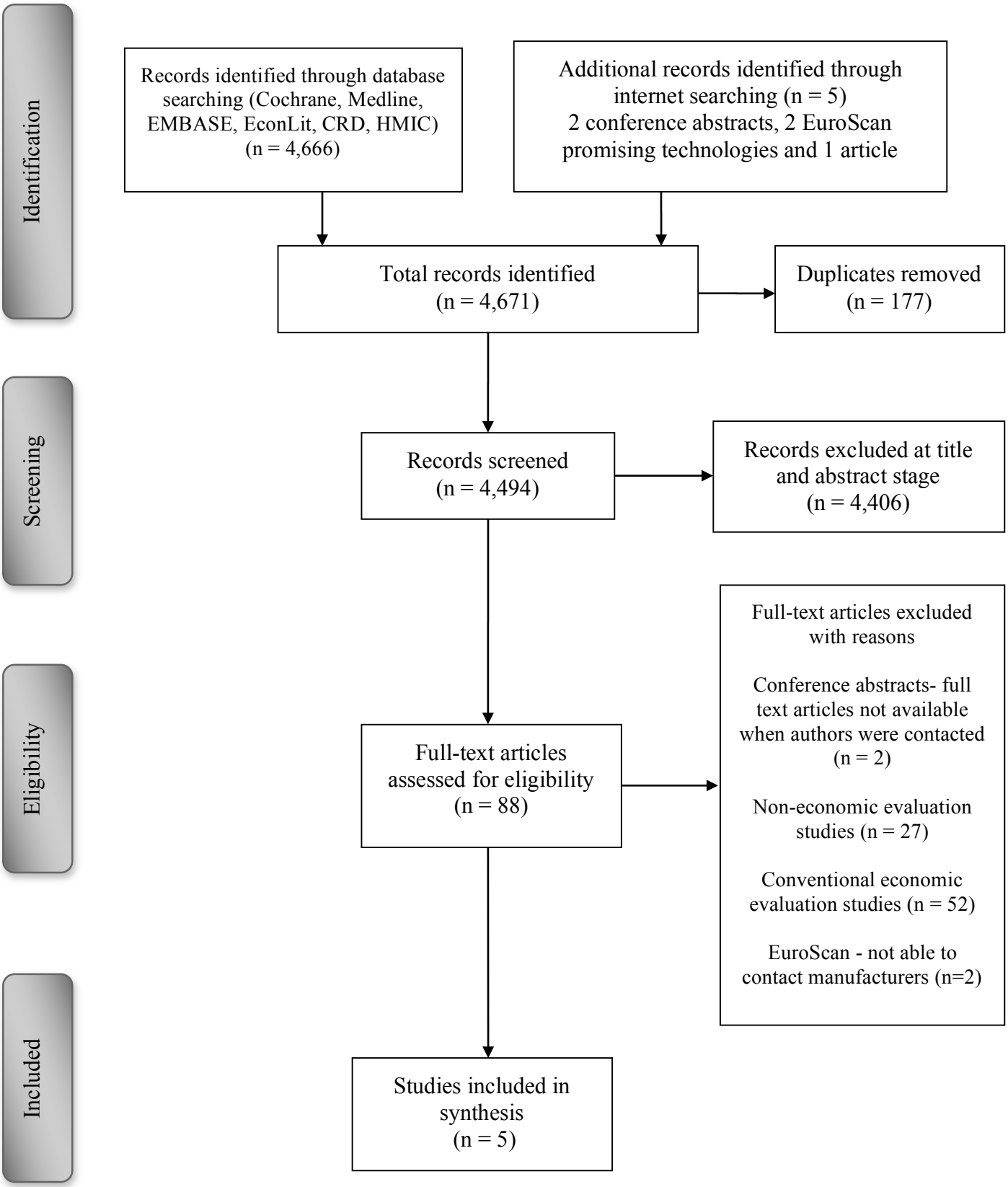


Table 3-3 Characteristics of included studies

Study	**Setting**	**Disease area**	**Study Population**	**Application of Test**	**Stage of development and evaluation**	**Index testing strategy**	**Comparator strategy**	**Modelling Approach**	**Type of economic analysis and Outcome measure**	**Perspective**	**Time Horizon**
Buisman et al (2016) Netherlands	Not stated	Rheumatoid arthritis	Inflammatory arthritis patients at risk of having rheumatoid arthritis	Diagnosis	Early development stage	B-cell related gene expression test IL-6 serum levels test MRI of hands and feet Genetic assays with susceptibility single nucleotide polymorphisms for rheumatoid arthritis	American College of Rheumatology/European League Against Rheumatism (ACR/EULAR) 2010 RA classification criteria	Decision tree followed by a Markov model	Cost-utility analysis Cost/QALY	Societal	5 years
Wu et al (2015) USA	Not stated	Persistent asthma	A hypothetical cohort of patients with persistent asthma ages 18-35 years	To predict response (ascertain whether a treatment has worked or not)	Early development	A pharmacogenomic test	Usual care (initiation of inhaled corticosteroids without testing)	Markov model	Cost-utility analysis Cost/QALY	Societal	10 years

Huang et al (2013) USA	STD clinical setting	Chlamydia trachomatis infection	Sexually active women attending an STD clinic ages ≥ 18 years	Screening	Early development	Experimental vaginal point-of-care test	Vaginal swab nucleic acid amplification tests (NAATs)	Decision tree	Cost-effectiveness analysis Cost /PID averted	Public Health care	Infertility (10yrs), ectopic pregnancy (5yrs), chronic pelvic pain (2years)
Vaidya et al (2014) Netherlands	Not stated	Peripheral artery disease	Patients with peripheral artery disease	To assess risk	Early development	The D-dimer strategy	Usual care (no risk assessment and uniform treatment	Markov model	Cost-utility analysis Cost/QALY	Societal	Lifetime
Bosch et al (2005) USA	Not stated	Coronary artery disease	Hypothetical cohort of 60-yr old male patients with coronary artery stenosis	Diagnosis	Concept stage (implicitly deduced)	Hypothetical catheter-based strategy	Usual care (patient with coronary artery stenosis present on angiography	Markov decision model	Cost-utility analysis Cost/QALY	Societal	Not stated

3.4.1 Study characteristics

The included studies, which were published between 2005 and 2016, are summarised in Table 3-3 above. Two studies were conducted in the Netherlands (Buisman et al., 2016; Vaidya et al., 2014) and three studies were conducted in the USA (Wu et al., 2015; Huang et al., 2013; Bosch et al., 2005). All studies stated the clinical condition explored, which spanned a range of disease areas, namely rheumatoid arthritis, persistent asthma, chlamydia trachomatis, peripheral artery disease, and coronary artery disease. All studies stated the intended application of the tests under evaluation. These were diagnosis, predicting a response to treatment, screening, and risk assessment. Four studies were conducted at the early development stage (Fig 2-3) (Buisman et al., 2016; Wu et al., 2015; Huang et al., 2013; Vaidya et al., 2014) and one study was conducted at the concept stage (Bosch et al., 2005). Model-based approaches were used in cost-utility and cost-effectiveness analyses. Cost-utility analyses were conducted in four studies (Buisman et al., 2016; Wu et al., 2015; Vaidya et al., 2015; Bosch et al., 2005), and a cost-effectiveness analysis was conducted in one study (Huang et al., 2013). The societal perspective was adopted in four studies (Buisman et al., 2016; Wu et al., 2015; Vaidya et al., 2014; Bosch et al., 2005) whereas the public healthcare perspective was adopted in one study (Huang et al., 2013). A time horizon of ≤ 10yrs was adopted in 3 studies (Buisman et al., 2016; Wu et al., 2015; Huang et al., 2013), a lifetime time horizon was adopted in one study (Vaidya et al., 2014), and one study did not state the time horizon adopted (Bosch et al., 2005).

3.4.2 Quality assessment of included studies

Based on the results from the quality assessment checklist, two studies had a score of 8.5 (Huang et al., 2013; Buisman et al., 2016), two studies had a score of 8 (Vaidya et al., 2014; Wu et al., 2015) and one study had a score of 7 (Bosch et al., 2005) (Table 3-4). All studies lost a point each because it was difficult to tell whether any important alternatives were omitted from the studies as insufficient information about clinical pathways were provided. Three studies lost a point each for not identifying and including all the important outcomes for each alternative (such as those from the patient's perspective) (Vaidya et al., 2014; Wu et al., 2015; Bosch et al., 2005). One study lost a point because it was not clear whether cost and consequences were valued credibly or adjusted for differential timing as all the data used were based on assumptions, but no information based on the assumptions was provided (Bosch et al., 2005). It is worth stating that due to word count restrictions, not all clinical information may be incorporated in the description of an economic evaluation (especially those conducted alongside randomised controlled trials). Usually these would have to be read together with other publications (such as clinical publications) to get an impression of the total evidence on a research question for evaluation. In this study, this was done before and during quality assessment to ensure an effective assessment.

Table 3-4 Summary of checklist scores

	Studies included in the review and score				
DRUMMOND 10-point checklist component	**Buisman et al (2016)**	**Wu et al (2015)**	**Huang et al (2013)**	**Vaidya et al (2014)**	**Bosch et al (2005)**
Was a well-defined question posed in answerable form	1	1	1	1	1
Was a comprehensive description of the competing alternatives given	0	0	0	0	0
Was the effectiveness of the programme or services established?	1	1	1	1	1
Were all the important and relevant costs and consequences for each alternative identified?	0.5	0	0.5	0	0
Were costs and consequences measured accurately in appropriate physical units?	1	1	1	1	1
Were the cost and consequences valued credibly?	1	1	1	1	0.5
Were costs and consequences adjusted for differential timing?	1	1	1	1	0.5
Was an incremental analysis of costs and consequences of alternatives performed?	1	1	1	1	1
Was allowance made for uncertainty in the estimates of costs and consequences?	1	1	1	1	1
Did the presentation and discussion of study results include all issues of concern to users?	1	1	1	1	1
TOTAL SCORE	8.5	8	8.5	8	7

3.4.3 Data extraction

Data was extracted for each study to answer the research questions and the results are presented in the ensuing sections.

3.4.3.1 How was the problem of insufficient data for model parameterisation managed?

To populate the index testing strategy arm of the models, several sources of data were utilized across the different studies from four main perspectives (test accuracy, costs, measures of effectiveness and transitional probabilities describing the disease states) and were found to be influenced by the stage of evaluation/analysis as shown in Table 3-5.

Table 3-5 Sources of data utilized in each study

Sources of data on index testing strategy				
Study	**Test accuracy**	**Costs**	**Measures of effectiveness**	**Transitional probabilities**
Buisman et al (2016) Netherlands	Estimates based on expert opinion (test developers), data was also from a recent evaluation of IL-6 serum level performance test and review of the literature	Estimate based on expert opinion (test developers), assumptions (basis of assumptions not stated) and, Dutch Healthcare Authority	REACH trial. Assumptions (based on literature and STIVEA trial), and the t-REACH trial	Elicited from the REACH cohort and DREAM cohort
Wu et al (2015) USA	Estimates based on assumptions (based on the literature on a different prototype pharmacogenomic test that has been developed)	Estimate based on expert opinion (test developers)	Published preference weights collected for the Asthma Policy Model via direct utility assessments using the time trade-off elicitation technique	Estimates based on assumptions (based on literature)
Huang et al (2013) USA	Primary data	Assumption - new test would be competitively priced and, therefore used the same point estimate for the comparator test	Primary data - downstream consequences of testing, and estimates based on literature	Primary data
Vaidya et al (2014) Netherlands	Assumption - perfect biomarker	MUMC-Maastrich University Medical Center	Published Literature	Published literature and estimates based on the REACH registry data
Bosch et al (2005) USA	Assumptions	Assumptions	Assumptions	Assumptions

One study was conducted at the concept stage (Bosch et al., 2005) and all the parameters used (test accuracy, costs, measures of effectiveness and transitional probabilities describing the disease states) were based on assumptions. However, no information was provided based on these assumptions. Sources of data utilized varied across the studies conducted at the early development stage (Table 3-5). It was noted that, in all these studies, the plausibility of the estimates utilized, and the robustness of the results obtained by utilizing these estimates were examined in sensitivity analysis.

3.4.3.2 What was the type of model used in modelling disease progression and the testing pathways (e.g. Decision tree, Markov)?

A Markov state transition model was utilized in three studies (Wu et al., 2015, Vaidya et al., 2014 and Bosch et al., 2005), a decision tree was utilized in one study (Huang et al., 2013) and both the decision tree and Markov model were utilized in one study (Buisman et al., 2016). In all cases, the Markov model described the disease and the decision tree described the testing pathways. Information was provided to justify the model structure in only three studies (Buisman et al., 2016; Wu et al., 2015; Vaidya et al., 2014).

3.4.3.3 Was test accuracy and all possible test results considered (i.e. true positives (TP), true negatives (TN), false negatives (FN), and false positives (FP))?

Each of the different possible test results were modelled explicitly in four studies (i.e., TP, TN, FP and FN) (Buisman et al., 2016, Wu et al., 2015, Huang et al., 2013 and Bosch et al., 2005). Vaidya et al (2014) assumed a perfect biomarker and therefore only considered TP and TN. Despite this assumption, no information was given on what the consequence of test error will be for this test. In one study, the first year of the 5-year time horizon was modelled as a decision

tree with chance nodes at 6 and 12 months as (repeated testing is part of the clinical pathway) to classify patients as TP, FP, TN and FN. Those classified as TP or FN at 12 months entered the patient level state transition model and followed for 4 years (Buisman et al., 2016). In another study, four subpopulations based on the test result of TP, FN, TN and FN were considered within the same model (Wu et al., 2015).

3.4.3.4 Were all the issues on the test-treat pathway that may be important to patients included in the analysis?

It is notable that, only one of the four studies that were conducted from the societal perspective acknowledged and included some issue on the test-treat pathway that may be relevant to the patient (Buisman et al., 2016). The issue of the patient's perspective was ignored in the other studies (Wu et al., 2015; Vaidya et al., 2014; Bosch et al., 2005). In one study, follow-up visit costs and productivity costs (measured by the number of days that a patient with a paid job was absent from work) were included (Buisman et al., 2016). In another study, the effects of how long patients were willing to wait for their test results was considered (Huang et al., 2013). This seemed reasonable as the study was conducted from a public health care perspective and one of the key mechanisms by which the test in this study might impact on outcomes was being able to treat patients at the time they are tested to prevent onward transmission of infection. Indeed, it was demonstrated in one-way sensitivity analyses in this study that one of the key parameters driving the results of the cost-effectiveness analysis was how long patients were willing to wait to obtain their results. The reason was that, a short processing time reduced the time between testing and treatment thereby increasing treatment rates and subsequent improvement in the quality of life of patients at the population level.

3.4.3.5 Did the analysis include headroom analysis, and at what stage of the analysis was it included?

Headroom analysis was noted to be included in the two studies conducted in the Netherlands (Buisman et al., 2016; Vaidya et al., 2014). It was also noted that in both these studies, neither was conducted at the concept stage of development but at the early development stage when there was some data available to describe the test parameters. Hence, headroom analysis was included in the early development stage but not the concept stage.

3.4.3.6 Was the uncertainty that accompanies early modelling acknowledged in sensitivity analysis?

The issue of a lack of data and the simplification of models to represent reality were acknowledged as study limitations in all the studies. It was noted that to deal with these limitations in all the studies, extensive sensitivity analyses was undertaken on all important parameters (e.g., sensitivity analysis of test accuracy) to evaluate the influence of uncertainty on model predictions. Probabilistic and deterministic sensitivity analyses were conducted in four studies (Buisman et al., 2016; Wu et al., 2015; Huang et al., 2013; Vaidya et al., 2014) and deterministic sensitivity analysis was conducted in one study (Bosch et al., 2005).

3.4.3.7 Was value of information (VOI) analysis conducted?

It was noted that VOI analysis was not conducted in any of the studies. However, considering the importance and role of VOI analysis in supporting decision making under conditions of uncertainty (typically characteristic of early economic evaluation), one would have expected it to be considered in the analysis.

3.5 Discussion

Typically, economic evaluations are conducted as a one-off exercise at the late stage of development of medical technologies (Vallejo-Torres et al., 2008). However, several studies have noted the importance of early economic analysis at the early stages of development of medical technologies to identify their potential economic value and likely impact, and to support and guide decision-making under conditions of high uncertainties (Sculpher and Buxton, 1997; Fenwick et al., 2006). However, there is little specific guidance on the implementation of early economic analysis. This systematic review focussed on exploring the methods and approaches that have been used in the early economic analysis of medical tests. Five studies were identified, and data extracted from each study to gain a greater understanding of how the problem of insufficient data for model parameterization has been dealt with, whether and how testing pathways have been modelled, and whether sensitivity analysis acknowledged the uncertainty that accompanies early modelling and the stage at which it is conducted.

The major issue associated with test evaluation is the rather difficult connection between testing and final health outcomes. However, in order to effectively evaluate the cost-effectiveness of a test, there is the need for the final health outcomes based on the downstream consequences of testing to be included in the analysis. This usually involves extensive modelling of delineated test-treat pathways. This is difficult enough to do for an established test, but for a new test, it is even more complicated and difficult to do because test-treat pathways may not be defined. This may explain why there are so few early economic evaluations of medical tests, and hence, the small number of studies identified in this study. Furthermore, early economic evaluations may take place within industry as a commercial in confidence activity whereby if disclosed, may result in damage to an organisation's

commercial interests, intellectual property or trade secrets, hence, the reason why such analysis may not be readily made available to the public.

To deal with the issue of an early analysis not having sufficient data for model parameterization, different sources of data were relied on by studies in this review and the sources of data used were influenced by the stage of the analysis (concept stage or early development stage). One study was conducted at the concept stage and as expected, all the data used were based on assumptions. However, no information on the basis for these assumptions were provided making these estimates somewhat arbitrary and weakly informed, although they were extensively examined in sensitivity analyses. Four studies were conducted at the early development stage and within the confines of this stage, varying data sources were utilized across the different studies. This observation is explained by the fact that the tests were at different phases of development even within the early development stage; thus, different levels of data were available specific to the different tests. For example, primary data was available for most of the parameters used in the evaluation in one study as the test was at an experimental phase. In the other studies, primary data were not available for all the parameters, thus, other plausible data sources were utilized to supplement primary data (making investigation of uncertainty particularly important). The plausibility and robustness of these estimates were however examined extensively in sensitivity analyses and their effects on the conclusions drawn from the models examined.

In all the studies, test accuracy was considered with four studies modelling explicitly each of the possible test results. This implied that, the full implications of test accuracy could be examined in the analysis. In the study that did not model each of the test results, this implied

that the consequences of test error were not examined. It was notable that even though a societal perspective was adopted in four studies, only one of these four considered and included in their analysis issues on the test-treat pathway that may be important to patients. Underpinning this statement is the fact that, the societal perspective is the broadest perspective to adopt in economic evaluations, and one would have expected that the issues on the testing pathway that are important to the patient would be more likely to be considered here compared to other perspectives (e.g., the hospital or healthcare system). This is important because including patient perspectives (issues that may be important to the patient on the test-treat pathway) could have a significant impact on the conclusions drawn from models. For example, in one study it was noted that how long patients were willing to wait for their test results was a key parameter driving the results of the cost-effectiveness analysis. Clearly, overlooking this important parameter in this particular analysis would have led to misleading conclusions. The most obvious and most studied way in which a test might produce benefits aside from improvements in test accuracy is probably the timing of test results (rapid tests). However, it is important to state that other aspects such as test acceptability to patients and professionals, procedural harms or benefits of the testing process are equally important and should also be considered (Ferrante Di Ruffano et al., 2012 and 2017).

Headroom analysis was included in two studies, and it was noted that in both studies it was not conducted at the concept stage but at the early development stage. However, the headroom method of analysis can be utilized as early as the concept stage to make a preliminary assessment of whether a test warrants further development. Then later in the early development stage when more data becomes available, this can be updated in the face of new evidence. This reduces the risk of investing in a technology that may not be economically viable. Though it may be argued that if a project is terminated at the early development stage not many resources

would have been invested, it is equally true that, if the headroom had been established at the concept stage and the project terminated at that stage, the invested resources could be used efficiently in another project or elsewhere. Thus, headroom analysis has the potential to promote efficiency at the beginning of the test development process.

Sensitivity analysis is an important undertaking in early economic analysis (which is characterized by a lack of data), as initial parameter estimates may have to be derived from data sources associated with high levels of uncertainty. All the studies included in this review acknowledged the uncertainty that is associated with the data used in their analysis and conducted extensive sensitivity analyses on important parameters to evaluate the influence of uncertainty on model predictions. It was also noted that VOI analysis was not conducted in any of the studies and this implies that no insights into the value of future research were obtained. Thus, there was the possibility of drawing wrong conclusions about whether to fund the further development of a new test or not based on the available evidence used. This observation could be explained by the fact that studies were conducted at a time when VOI analysis was not well established as a concept or it was deemed irrelevant.

3.6 Strengths and limitations of the study

To the best of our knowledge, this is the first study that has focussed on systematically exploring the tools and methodologies that have been employed in the implementation of early economic analysis of medical tests. It has contributed to gaining a greater understanding of what currently pertain in this field and has identified areas requiring further research in order to maximise the value of such analysis.

Like every study, this study has its own limitations. Few studies were found to be eligible to be included in the review. Having additional studies from the industry would have supplemented the review and given a broader and more holistic view of the current state of the implementation of early economic analysis of medical tests. Furthermore, it was not possible to compare results as studies were conducted to answer different research questions.

3.7 Recommendations

1. To fully capture the potential benefits of testing on patient relevant outcomes and thus the potential health economic impact of medical tests, the assessment of the outcomes of testing should transcend health and specific payer perspectives to include all the issues on the test-treat pathway that may be important to the patient. This will mitigate against over- or underestimating the true value of tests in early economic analysis and thereby appropriately informing decision-making.
2. The full implications of test accuracy should be considered in an analysis by including all the possible test results and their subsequent patient pathways in a model.
3. The influence of uncertainty on model predictions should be evaluated through extensive sensitivity analyses of all the important parameters.
4. The potential adoption of VOI analysis should be considered in early economic analysis studies. This can be beneficial in mitigation against drawing incorrect conclusions about whether to fund the further development of a test.
5. Headroom analysis has the potential to provide information about the viability of developing a new test and should be considered at the early stages of development, particularly at the concept stage to promote efficiency at the start of the test development process.

3.8 Suggestions for future work

The review has shown that some of the methods available (VOI and Headroom analysis) are not being utilized in the early economic evaluation of test. Is there a need to refine or develop these methods in this specific context? To answer this question, further research is needed in this field.

3.9 Conclusion

Having provided a greater understanding of how economic evaluations focussed on medical tests have been conducted at the early phases of development, the next chapter (Chapter 4), is focused on delineating the existing test-treat pathway(s) for typhoid fever in Ghana. This was undertaken to provide the necessary background information to undertake an early economic evaluation of a typhoid test, which is described later in this thesis.

Chapter 4: Delineating the test-treat pathway(s) for typhoid fever in Ghana: a case study using the framework approach

This chapter is focussed on exploring and delineating the existing test-treat pathway(s) for typhoid fever in Ghana. Using the framework approach, it highlights the potential role of the application of qualitative methodologies in enhancing the process of test-treat pathway delineation. This was necessary so that the models developed for the early economic analysis (see Chapter 6) reflected the current practice with respect to typhoid testing in the Ghanaian setting.

4.1 Background

For an informative economic model to represent appropriately the disease, persons not having the disease, and misclassifications, and for the model to be fit for purpose for decision making, an in-depth understanding of the care pathway is required (Briggs et al., 2006). Therefore, delineating care pathways and any existing variations, is a key information requirement to effective medical test evaluation (see Chapter 2, p.18-20).

Delineating the care pathway is a difficult undertaking, and qualitative methods can play a key role in enhancing this process (Chilcott et al., 2010; Kaltenhaler et al., 2012; Roberts et al., 2012; Kaltenhaler et al., 2014; Husbands, 2016). For example, qualitative methods can serve as a medium through which the perspectives of all the different types of experts knowledgeable on a topic area can be captured (Kelley et al., 2003). These may include capturing variability in practice when defining care pathways to assist in model structure conceptualisation (Iglesias et al., 2016). Variability may arise in different ways in practice. For example, it may arise from clinicians and patients involved at different points in a care pathway and in different settings

or locations. Capturing variability can potentially assist in improving the validity as well as the generalizability and credibility of models. Different qualitative research methods and the studies in which they have been applied to the model process include the Delphi technique (Sullivan and Payne, 2011; Iglesias et al., 2016), focus groups (Roberts et al., 2012), and stakeholder workshops (Squires et al., 2016). Despite the potential usefulness of qualitative methods, there has been limited uptake of formal qualitative methods in model development by modellers and health economists (Husbands et al., 2017). Possible explanations for their limited uptake are the general assumption that qualitative methods are resource intensive, and the lack of familiarity of their potential to assist with model development.

Information on the care pathways for typhoid fever in Ghana, their variations, diagnostic challenges (which are likely to arise from the different levels of care) and the potential placement and role (replacement, triage or add-on) of a new test on existing care pathways was lacking, and research was needed to be done to inform it.

4.2 Study aims and objectives

The primary aim of this chapter was to gain insight into the current practice (to delineate existing care pathways) of typhoid fever testing and treatment in Ghana with a view to using the information for later model structure conceptualisation. The secondary aim was to highlight how a qualitative approach can be applied in a real-world context to gain insights into current practice. Specifically, the work reported in this chapter sought to:

1. Describe the existing care pathway for typhoid fever in Ghana and any variations.
2. Explore the reasons for any variations in the existing care pathways for typhoid fever in Ghana.

3. Capture the views of a range of healthcare professionals working in Ghana on the potential role of a hypothetical new rapid test for typhoid on the care pathway they work in.

4.3 Methods

The perspectives of a range of healthcare professionals working in different settings and across different practices in the Eastern region of Ghana were explored to answer a series of questions: What is the test-treat pathway for typhoid fever in Ghana? Are there variations in the existing test-treat pathway? Finally, what would be the potential role(s) of a new test on the existing pathway in Ghana? One region was chosen as a case study because the literature available on the structure of the health care delivery system in Ghana indicated that, variations in practice were likely to result from the different levels of care (which is the same across all regions) rather than resulting from the individual regions. The Eastern region was selected because the Principal Investigator (PI) was familiar with the region having previously lived and worked there. Thus, it was comparatively less resource-intensive for the PI to collect the necessary data from this location. A purposive, maximum variation sampling approach was adopted. This involved approaching all those who were knowledgeable on the topic of interest and could give potentially different perspectives on the topic (Kelley et al., 2003). The results of the study were compared between health care settings because this was the main factor likely to drive any variations in practice.

4.3.1 Participants

To be eligible for inclusion in the study, the participant had to meet the inclusion criteria of being a clinician or medical assistant who worked in a healthcare provision facility (Health Centre, Polyclinic, hospitals, etc.) in the Eastern region of Ghana. These experts were chosen

to participate in this study because they were considered knowledgeable in the topic of interest and thus were in a better position to provide valuable information to aid the study. To avoid sampling bias, participants of all ages and genders were included.

4.3.2 Data collection (March-July 2017)

Some level of pragmatism was required to choose the most appropriate approach to collect data in this study because, in-depth interviews were not possible due to participants' time constraints. Consequently, a self-administered survey using open-ended questions was used (Jansen, 2010). Acknowledging the limitations of using a survey approach (i.e., responses were likely to lack depth, and a survey would not allow probing by the researcher or checking of the participants' understanding of the questions), a well written questionnaire was required to elicit adequate valuable information. A questionnaire was piloted in the Greater Accra region and revised based on responses into the final version used in this study. After ethics approval from the University of Birmingham (ERN_16-0947) and from the Ghana Health Service Ethical Review Committee (GHS-ERC: 03/12/16), medical superintendents of various health facilities were contacted to arrange a meeting between the researcher and the clinicians and/or medical assistants working at their facilities. At these meetings, the aims and the expected outcomes of the study were explicitly explained to the experts and interested persons were recruited into the study after having gone through the informed consent process and signed the consent form. In each institution, groups of recruited participants were given the opportunity to go through the questionnaire. This was done to ensure that they understood the questions, and explanations were offered to clarify issues where necessary. To enable easy analysis of the data set, questionnaires were coded for easy identification of the facility and its location within the region. Participants were anonymized. A total of 70 self-administered questionnaires were distributed to the participants by the researcher in person.

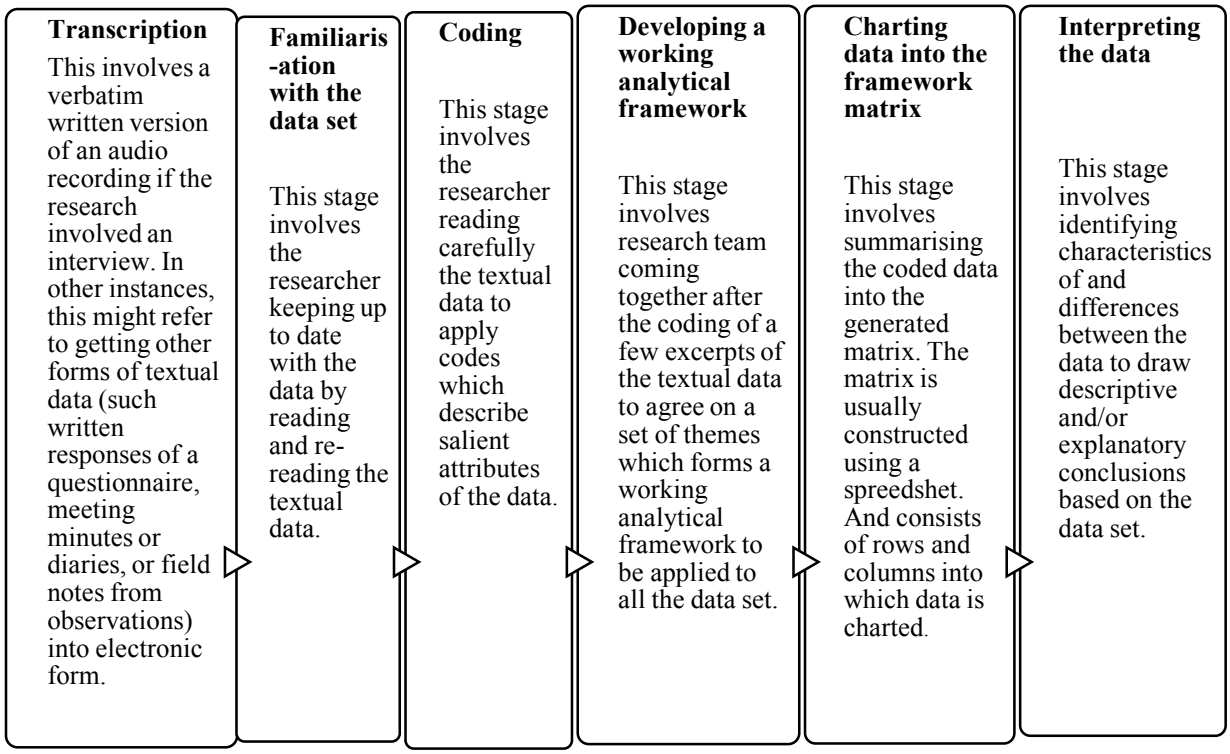
The following themes were generated *a priori* and informed the content of the questionnaire: presumptive treatment views (treatment of clinically suspected cases without, or prior to, confirmatory laboratory test), diagnostic test(s) used, test negative management (i.e., care plan for those with a negative test result), test positive management (i.e., care plan for those with a positive test result), prescribed antimicrobial for test positives, dosage of antimicrobial, assessment of test-treat outcome and intended role of new test. All the themes except "intended role of a new test" were generated to gain insight into current practice. To do this, these themes had to be considered together. Thus, they are reported and discussed together under the section "delineating existing test-treat pathway and reasons for variations". The theme "intended role of new test" was generated to capture the views of the participants on the potential role of a hypothetical new test, and where it should be placed on the testing pathway to improve current practice. To capture these views, participants were asked to choose a statement which best described the role of a hypothetical new rapid diagnostic test for typhoid to them: (a) to test patients first to decide on who should receive further testing with the test you use now (triage); (b) as the main investigative tool without further testing with the test you use now (replacement); or (c) to provide additional diagnostic information after testing with the test you use now (add-on). The questionnaire is shown in Appendix 2.

4.3.3 Data analysis

The analytical framework adopted for analysing the data collected in this study is framework analysis (Gale et al., 2013). Framework analysis belongs to a wider group of analysis methods (known as thematic analysis) that seeks to identify, analyse and report patterns within data (Braun and Clarke, 2006). The data collected using this method are based on pre-defined themes (deductive) rather than emerging from data (inductive). The matrix output is the distinctive attribute of this method of analysis (Pope and May, 2006). It provides a systematic

structure into which data is charted (inputted) to enable in-depth analysis. The matrix consists of rows, columns and cells of summarised data. In this case study, rows referred to participants, columns referred to the pre-defined themes, and cells of summarised data referred to participants' responses. Framework analysis was chosen in this study particularly due to its deductive approach, which makes it especially suited to research that has specific data needs set in advance (Srivastava and Thompson, 2009). Adopting this method of analysis in this study allowed for pre-defined themes to be developed that shaped the nature of the data collection. Furthermore, framework analysis was chosen because it is intuitive and allows for easy recognition of patterns in the data set once a matrix has been developed (Gale et al., 2013). In addition, the systematic approach embedded in the framework method of analysis ensures that the rigour of the analytical process is maintained, thereby enhancing the credibility of the findings (Ritchie and Lewis, 2003). Figure 4-1 illustrates the procedure for analysis using the framework approach.

Fig 4-1 Procedure for analysis using the framework approach



Written responses retrieved from participants were first converted into electronic form and read repeatedly to ensure familiarisation with the data. This provided a good opportunity to become immersed in the data set and have a better recollection of key information (Braun and Clarke, 2006). To classify all the data set according to the pre-defined themes, the textual data was colour coded. There were no emerging new themes from the survey; thus, the pre-defined themes generated were then developed into a working analytic framework, a matrix consisting of rows and columns into which data was charted to enable in-depth analysis (Pope and Mays, 2006). The "analytical framework matrix" is shown in Appendix 3.

4.4 Results

The findings of the study are reported under the following sections: demographics of participants, delineating existing test-treat pathway(s) and reasons for variations, and the perceived role of a new test.

4.4.1 Demographics of participants

Fifty-one questionnaires were retrieved in total representing a 73% response rate. All non-respondents were noted to be participants working in primary hospital facilities, and time constraint was the main reason given for non-response. Twenty responses were from participants recruited from secondary care settings and thirty-one were from participants recruited from primary care settings. To ensure participant anonymity, the names, gender and age of the participants were not captured.

4.4.2 Delineating existing test-treat pathway and reasons for variations

Two main pathways were delineated from the analysis of the matrix: The Widal care pathway (suspected typhoid cases have Widal test) (Fig 4-2) and the culture pathway (suspected typhoid cases have culture, i.e., blood stool or urine) (Fig 4-3). Underpinning this variation was the type of health care facility from which the participants were recruited. It was noted from the analysis of the matrix that, the Widal care pathway was predominantly adopted in primary health care facilities (typically district level facilities that provide the most basic care). And the culture pathway was predominantly adopted in secondary hospital facilities (typically regional hospitals which are an upgrade of primary hospitals in terms of infrastructure and resources). There were two exceptions to these observations. **Participant 50**(P) stated blood or stool culture as the diagnostic tool used even though the participant was recruited from a primary hospital facility and **Participant 31**(S) stated that the Widal test was used even though the participant was recruited from a secondary hospital facility (letters P and S in parenthesis indicate whether a participant was recruited from primary care or secondary care respectively). **Participant 31**(S) stated, *"it is the commonest and simple to do"* as the reason for using the Widal test. This quote is suggestive of the fact that the Widal test could be available in secondary hospitals even though culture might be the preferred test in such settings. **Participant 50**(P) **stated**, *"it is readily available in our lab"* as the reason for employing culture.

It was noted that, the management aspect of the two care pathways following testing with either the Widal test or culture was similar. Test negative patients were investigated further for other conditions especially malaria and treated according to the final diagnosis in both pathways. For example, **Participant 16**(P) stated, *"when negative, we check for other conditions like malaria*

and treat them". **Participant 19**(P) stated, *"I look out for other conditions presenting like typhoid fever in typhoid-negative clients and treat"*. **Participant 35**(S) stated, *"when negative I investigate the underlying cause of symptoms and treat appropriately"*. And **Participant 38**(S) stated, *"when negative, I investigate further to get the specific condition"*. From the analysis, it was noted that typhoid positive patients were treated with antibiotics, mainly oral ciprofloxacin at a dose of 500mg every 12h for a period of 7-14 days. After a course of treatment, they were reviewed, and in most cases re-tested to ascertain whether the test-treat process was successful in improving the patient's health. For example, **Participant 19**(P) stated, *"I treat typhoid positive patients with oral ciprofloxacin. Dosage: Tb Ciprofloxacin 500 mg bd* 7-14 days"*. **Participant 25**(P) stated, *"positives: I treat with ciprofloxacin. Dosage: Tb Ciprofloxacin 500mg bd *14 days"*. And **Participant 32**(S) stated, *"I give oral antibiotics medication for positive patients. Dosage: Tb Ciprofloxacin 500 mg bd* 10 days"*. It was noted that the choice of oral ciprofloxacin was mainly due to it being the first-line treatment of choice in typhoid fever management and also for other reasons such as its efficacy against the causative organism, availability and cost. For example, **Participants 31**(S) **and 51**(P) stated, *"it is the first-line drug to manage typhoid",* and **Participant 33**(S) stated, "*because of its efficacy and sensitivity against the causative organism Salmonella typhi*". In addition to this observation, some participants prescribed alternative antibiotics in addition to oral ciprofloxacin. For example, **Participant 1**(P) stated, *"positive cases are put on treatment. Dosage: Tb Metronidazole 200-400mg tds* 7days"*. **Participant 4**(P) stated, "*treat positives with antimicrobials. Dosage: Tb Cefixime 200mg bd*7 days"*. **Participant 8**(P) stated, *"I give antibiotics for test positive cases. Dosage: Tb Azithromycin 1g daily* 7days"*. And **Participant 30**(S) stated, *"positives: I give antibiotics. Dosage: Tb Cefixime 400mg bd*10-14 days"*.

Of further interest in this study was to explore the views of participants on applying presumptive treatment for typhoid fever. It was noted that, the majority of participants were averse to presumptive treatment, because they expressed the opinion that this may lead to treatment failure and the development of antimicrobial resistance. For example, **Participant 18**(P) stated, *"it could lead to treatment failure and antimicrobial resistance"*. **Participant 28**(S) stated, *"it can cause antibiotic resistance"*. **Participant 11**(P) stated, *"presumptive treatment may be prophylactic but can cause antimicrobial resistance"*. And **Participant 1**(P) stated, *"typhoid fever presents like malaria and presumptive treatment may lead to treatment failure"*. However, it was also noted that few participants advocated for presumptive treatment, but only under certain circumstances. For example, **Participant 50**(P) stated, *"since culture usually takes 72 h, presumptive treatment can be started"*. **Participant 30**(S) stated, *"this is a good intervention to prevent complications since the most precise blood/stool cultures are not available in most district and health centres"*. And **Participant 23**(P) stated, *"Widal test is not wholly reliable hence treatment can be made if test is not done but clinical features are present"*. It is clear from these quotes that; presumptive treatment appears to be used because of delays in diagnosis or concerns about the accuracy of the Widal tests or lack of availability of a better test.

Fig 4-2 The Widal care pathway

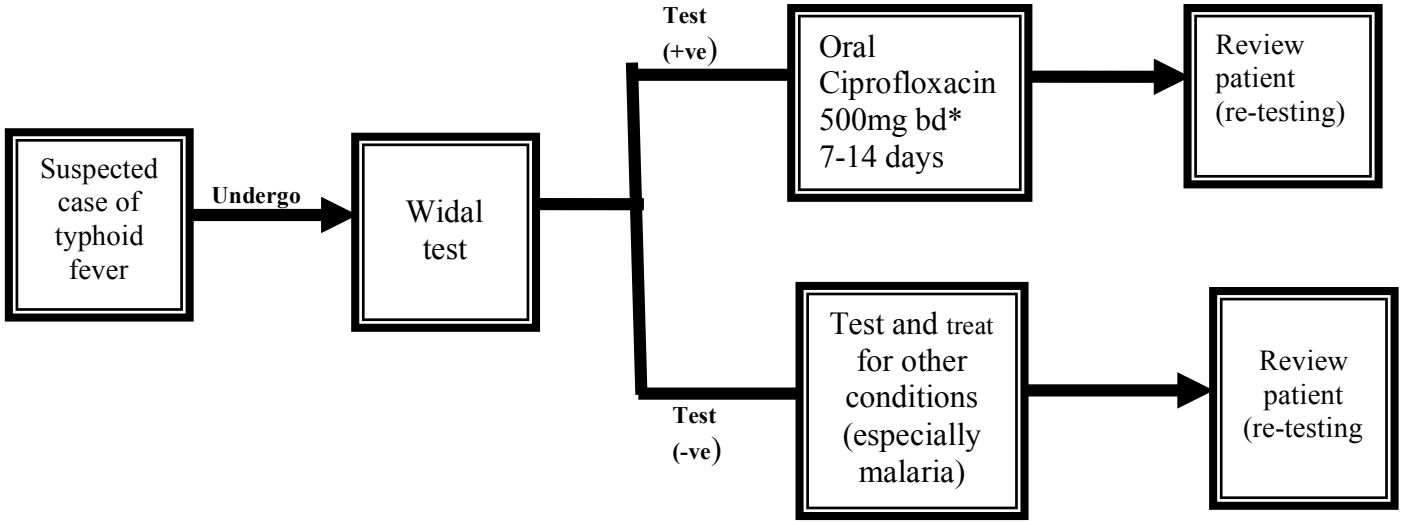
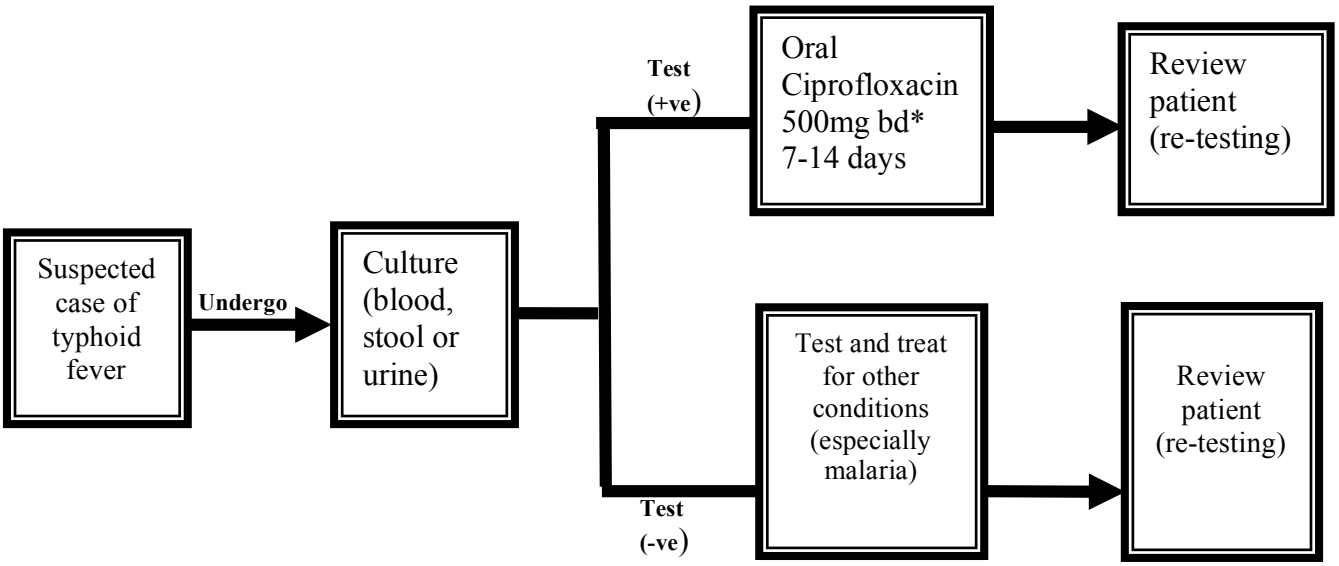


Fig 4-3 The culture care pathway



4.4.3 Perceived role of new test

Twenty-eight participants from across the different levels of care stated that they saw themselves using a new test, *"to provide additional diagnostic information after testing with the test they were using currently (add-on)"*. Sixteen participants from across the different levels of care stated that they saw themselves using a new test, *"to test patients first to decide on who should receive further testing with the test you use now (triage)"*. And six participants from across the different levels of care stated that they saw themselves using a new test, *"as the main investigative tool without further testing with the test they were using currently (replacement)"*. This observation was further explored according to health care setting, and a similar trend in the response was noted by health care setting.

In addition to exploring the participants' views on the role of a new test, their views on what test properties the new test should possess were also explored because, the chosen role of a test implies the necessary properties it should possess (see Chapter 2, p.20-22). It was noted that, most of the participants (32) wanted the new test to be more sensitive followed by those who wanted it to be more specific (15). **Participants 8**(P) and **30**(S) stated that, they wanted the new test to be more sensitive and more specific compared to the test they were currently using. **Participants 13** (P) and **24** (P) did not answer the question. The following are some other properties as stated by the participants: improved reliability, low cost, greater availability, better accuracy, the ability to use other fluids apart from blood, easy to use, fast results, and high predictive values. For some participants, it was noted that their views on what test properties a new test should have did not correspond with their choice of what the role of the test should be. For example, **Participant 37**(S) stated *"better accuracy, low cost, accessibility and ease to perform"* as the properties of the new test but wanted the test to be used as an add-on to culture. However, it is intuitive that from these stated properties, the test should ideally

be placed before culture in the care pathway to enable expedite clinical decision- making to ensure prompt implementation of an appropriate effective treatment.

4.5 Discussion

An in-depth understanding of test-treat pathways is key to the effective economic evaluation of medical tests. Qualitative methods can play a potentially key role in enhancing this process despite their limited uptake in practice. This study focussed on gaining insight into current practice for typhoid fever testing and treatment in Ghana by utilizing the framework analysis method. Two main care pathways for typhoid fever in Ghana were identified (Widal and culture pathways), and the majority of the participants were averse to applying presumptive treatment to typhoid fever and saw themselves using a new rapid test if it was introduced into clinical practice as an add-on test.

The two main pathways identified in this study were primarily due to variations in the laboratory diagnostic capacity of healthcare facilities from which the participants were recruited. The Widal test is a simpler technology to employ compared to culture and therefore was more likely to be employed in primary hospitals which lacked resources, whereas culture was more likely to be employed in secondary hospitals. It was noted that, although most participants were averse to applying presumptive treatment to typhoid fever, a few advocated for presumptive treatment. They did so for valid reasons such as delays in diagnosis or concerns about accuracy of test or lack of availability. These circumstances reflect the characteristics of current practice in Ghana and therefore a new test would need to be "better" in these respects for healthcare professionals in Ghana to use if current practice is to be improved. The following were stated by participants as the properties that a new test should have to improve current practice: improved reliability, low cost, greater availability, better accuracy, the ability to use

other fluids apart from blood, easy to use, fast results, and highly predictive. These findings provide key information that cannot be ignored in any effort to develop a new test to be used in the Ghanaian setting. The use of alternative antibiotics in addition to ciprofloxacin as noted from this study demonstrates the potential for waste in current practice, indicating the need for standardisation in, or adherence to, the treatment regimen for typhoid fever patients.

It was noted that regardless of the setting, most participants saw themselves using a new rapid test if it was introduced into clinical practice as an add-on test. However, one would have expected responses to vary depending on the setting in which they worked. This viewpoint is espoused on the premise that, if the main limitation of the Widal test as acknowledged by the participants who follow the "Widal care pathway" is unreliability of the test, then it is intuitive that a majority of them should be more inclined towards using a new test which is intended to be more reliable as their main investigative tool rather than as an add-on to the test they use currently. The argument might be that the very limitation of the Widal test is the reason why the new test should be used as an add-on test to the Widal test (to provide additional diagnostic information). However, the counterargument is that if the new test is accurate enough to be used to confirm diagnosis after the Widal test, then it is intuitive that it will replace it to save money (although it is acknowledged that such a decision should be informed by a cost-effectiveness analysis comparing the two tests). Also, it was expected that for those participants who follow the "culture pathway", most of them would be more inclined towards using the new test to triage patients in their setting. Underpinning this argument is that, if culture takes 2-3 days (causing potential treatment delays), a triage test may be used to rule out diagnosis, so triage test negatives are discharged, and triage test positives receive further testing to confirm the diagnosis. However, in this study it was not possible to explore the reasoning behind the participants' choices (which could be due to a lack of understanding, perception of

the accuracy of the existing tests, scepticism about the reality of a new test, acceptability and accessibility of a new test).

Furthermore, it was noted that there seemed to be a lack of understanding by health professionals (clinicians and medical assistants) in Ghana on test characteristics. The clinical implications are that, clinicians and medical assistants in Ghana are likely to order more laboratory tests than required, order inappropriate tests which will consequently result in inappropriate treatment being administered, all of which have obvious implications for the quality of patient care. Also, there are large socioeconomic implications as inappropriate or superfluous tests may significantly add up to healthcare expenditure. This highlights the need for further education for clinicians and medical assistants in Ghana on the diagnostic properties of available test to improve the quality of patient care whilst decreasing cost to the healthcare system.

The findings from this survey informed the modelling process in Chapter 6 in a number of ways. First, the delineated test-treat pathways informed and assisted in defining the boundaries and structures of the models utilized. The findings also indicated that, the settings in Ghana are different depending on the level of care, and its implication was that any cost-effectiveness analysis should consider these alternative pathways separately (and this was duly followed). Furthermore, the findings of this research assisted in identifying the appropriate comparator testing strategies required for the cost-effectiveness analysis. Also, in order to have effectively and appropriately evaluated the true value of the new test, its intended role in the care pathway was to be clearly defined in the model to enable the assessment of the test in that role. Although, the study findings indicated that there was no consensus amongst participants on what this role should be, this finding highlighted the need to evaluate the new test in all the possible roles in

each of the delineated pathways.

Adopting the framework method worked well in this study. The relatively well structured and deductive approach to questioning enhanced the efficiency of the data collection process because it allowed for the information requirements to be well specified in advance which shaped the nature of the data to be collected. Furthermore, the matrix output allowed patterns to be more easily identified. For example, identifying the test-treat pathways once the matrix had been developed and data charted became far easier. Using an open-ended approach to questioning also allowed the identification of unanticipated issues. An example is the observation of the use of alternative antibiotics in addition to the first-line treatment of choice. Also, it was possible to capture and analyse the perspectives of 51 participants working in different settings or locations and across different practices (rather than a few informants) to gain an in-depth understanding of current practice by adopting the framework approach.

Furthermore, due to time and resource (money and skill) constraints, many techniques that might be used to elicit test-treat pathways are not feasible. For example, there are challenges such as the practicality of organising and running face-to-face discussions in focus groups. There is also the challenge of gaining the inputs of multiple clinicians with busy schedules by asking them to complete questionnaires in a series of rounds as in the Delphi approach. As has been shown in this study, framework analysis is an example of a qualitative methodology that is likely to be more accessible and feasible across a wide range of health economic settings. In principle, the framework method of analysis can be used to analyse different forms of textual data including interview transcripts, responses to a questionnaire, meeting minutes and field notes, to produce highly structured summarised data across a wide range of research areas (Pope et al., 2000). The appeal of the framework method of analysis to quantitative researchers

is that, it is deductive and allows qualitative information to be collated within pre-defined categories or themes by the researcher. Thus, for this method to be beneficial in any health economic context, it is important for the researcher to be clear from the outset the issues to be explored to inform the most appropriate questions to be asked during data collection. The researcher should be clear from the outset on issues such as the purpose of the model, the type of model to be developed and the elements that are heterogeneous. Once data is collected and converted into textual form, framework analysis can then be used to analyse the data set to generate outputs that may appropriately inform the model development process. It is worth stating that in qualitative research when reporting findings, the researcher usually describes and interprets quotes as without context they are not self-explanatory. However, in this study, because a survey rather than interviews was conducted, there was limited interpretation: this was due to the nature of questions asked.

Clearly, the modelling process can be improved by the potential application of qualitative methods and those working in this field should consider the opportunities they provide. Qualitative investigation is a step beyond the informal discussion that modellers normally have with clinicians to inform the model development process.

4.6 Strengths and limitations of the study

The strength of this study is that, it highlights how qualitative methods can potentially assist in improving the model development process by facilitating the involvement of, and capturing and analysing the perspectives of the different types of experts knowledgeable on a topic of interest and working in different settings (rather than a few informants), thereby potentially enhancing the credibility and generalizability of models developed. Furthermore, the approach adopted in this study is intuitive and deductive in nature (allowing qualitative information to

be collated within pre-defined themes), thus, can be readily adopted by quantitative researchers.

The limitations of this study include the fact that only clinicians and medical assistants were sampled. If other relevant stakeholders such as patients (van Voorn et al., 2016) and policy makers had been sampled, they may have held different views or added further considerations. Furthermore, it was not clear whether the participants understood all the questions and it was not possible to probe their responses because the interviews were not personally conducted. For example, in this study, this was a deficiency when capturing valid participants' views about the anticipated role of a new test.

4.7 Conclusion

Having gained a greater understanding of the intricacies of the typhoid test-treat pathway(s) in Ghana, the next chapter (Chapter 5) is focussed on exploring the literature on economic evaluations of typhoid. This was to acquire knowledge specific to typhoid economic analysis to augment the knowledge already acquired to ensure that the most appropriate methodologies are utilized in Chapter 6 in the evaluation of the potential cost-effectiveness of the hypothetical test.

Chapter 5: Economic evaluation of typhoid - a review

This chapter is focussed on exploring the literature on the economic evaluation of typhoid. The focus on typhoid economic evaluation in general was to gain a holistic understanding of the modelling approaches that have been adopted in this setting for the economic evaluation of typhoid interventions, and the circumstances under which they were adopted and utilized. The findings of this chapter can then contribute to informing the modelling process utilized in Chapter 6.

5.1 Background

The modelling approach adopted and utilized in economic evaluations is informed by the natural history of the disease, care pathway(s) and the type of intervention being evaluated (Brennan et al., 2006). Two main types of interventions exist for typhoid fever. These are interventions targeted at typhoid treatment (such as test treat strategies), and interventions targeted at typhoid prevention (such as vaccination) (WHO, 2003). The different nature of these interventions imply that they will require different modelling approaches in their evaluation.

Model-based cost-effectiveness studies involve adapting an existing model or developing a new model for the analysis. Identifying what has already been done in the area of interest is fundamental to the approach taken. Thus, in order to examine the value of the hypothetical test for typhoid fever in Ghana, there was the need for a review of previous typhoid economic evaluations to understand how the different interventions for typhoid fever have been modelled.

5.2 Study aim

The aim of the review was to explore the literature focussed on typhoid economic evaluations to gain insights into the types of models that have previously been adopted in this setting (with particular interest in test-treat evaluations), and to capture data on model inputs that may be useful for a *de novo* model.

5.3 Methods

For this review, the following databases were searched for studies published from inception to September 2017. No restrictions on language were applied.

1. Medline
2. Excerpta Medica dataBASE (EMBASE)
3. Centre for Reviews and Dissemination [Database of Abstracts of Reviews of Effects (DARE), Heath Technology Assessment (HTA), and NHS Economic Evaluation Database (NHS EED)]
4. PubMed

A reference search was undertaken by scanning the bibliographies of the relevant articles finally included in the review for any additional relevant articles. The reference management software Endnote was used to manage the list of articles.

5.3.1 Search terms

The search strategy was customized for each database and searching was undertaken using the following terms, including truncation of terms where appropriate: economic evaluation, cost-benefit analysis (CBA), cost-effectiveness analysis (CEA), cost-utility analysis (CUA), typhoid fever and enteric fever. The search strategy was developed to cover all types and forms

of economic evaluations to ensure that all the relevant studies were retrieved first before focussing on model-based studies. This approach was taken because a scoping search undertaken showed that there was not much literature in this field, and the intention was to keep the search broad enough not to miss out on any study due to the limited literature base. Appendix 4 shows the complete search strategy for each database.

5.3.2 Inclusion criteria

Studies were deemed potentially relevant to be included in the review if they were:

1. Economic evaluations focussed on typhoid fever; or
2. Systematic reviews of typhoid economic evaluations.

5.3.3 Exclusion criteria

Studies were excluded if they were:

1. Not in English; or
2. Not conducted in an endemic setting; or
3. Trial protocols or commentaries; or
4. Letters or editorials.

5.3.4 Selection of articles for the review

After the removal of duplicates, a two-stage approach to article selection was used. The first stage involved screening of abstracts and titles to identify potentially relevant articles whilst the second stage involved screening of the full text articles. The screening process during both stages was conducted against the inclusion and exclusion criteria. The screening process was

undertaken by two reviewers (S.F and P.B) independently and any disagreements regarding eligibility were resolved using the opinion of a third reviewer (C.D) where necessary. All studies identified after the second stage of article selection were included for data extraction.

5.3.5 Data extraction

A data extraction form was developed in Excel and piloted. Then data extraction was conducted for each study included in this review to answer the following questions:

1. What were the interventions evaluated (test-treat strategies; vaccination)?
2. What was the economic evaluation approach adopted (CUA or CBA or CEA) and what was the outcome measure?
3. What type of model was used (static model; transmission dynamic model)?
4. What was the impact of the intervention on the transmission of infection between individuals?

5.3.6 Quality assessment

The methodological quality of the included studies was assessed using a modelling quality checklist modified from Philips et al. (2006). This modelling quality checklist was adopted because, although the search strategies were developed to capture the different forms of economic evaluation, the focus of this study was to examine how typhoid interventions have been modelled. Thus, a modelling quality checklist was considered appropriate here. Each item on the checklist was rated under the categories “yes”, “no”, “unclear” and “not applicable” by the extent of reporting. No study was rejected on quality grounds because the focus of this review was to explore the models used and not to comment on the validity of results and

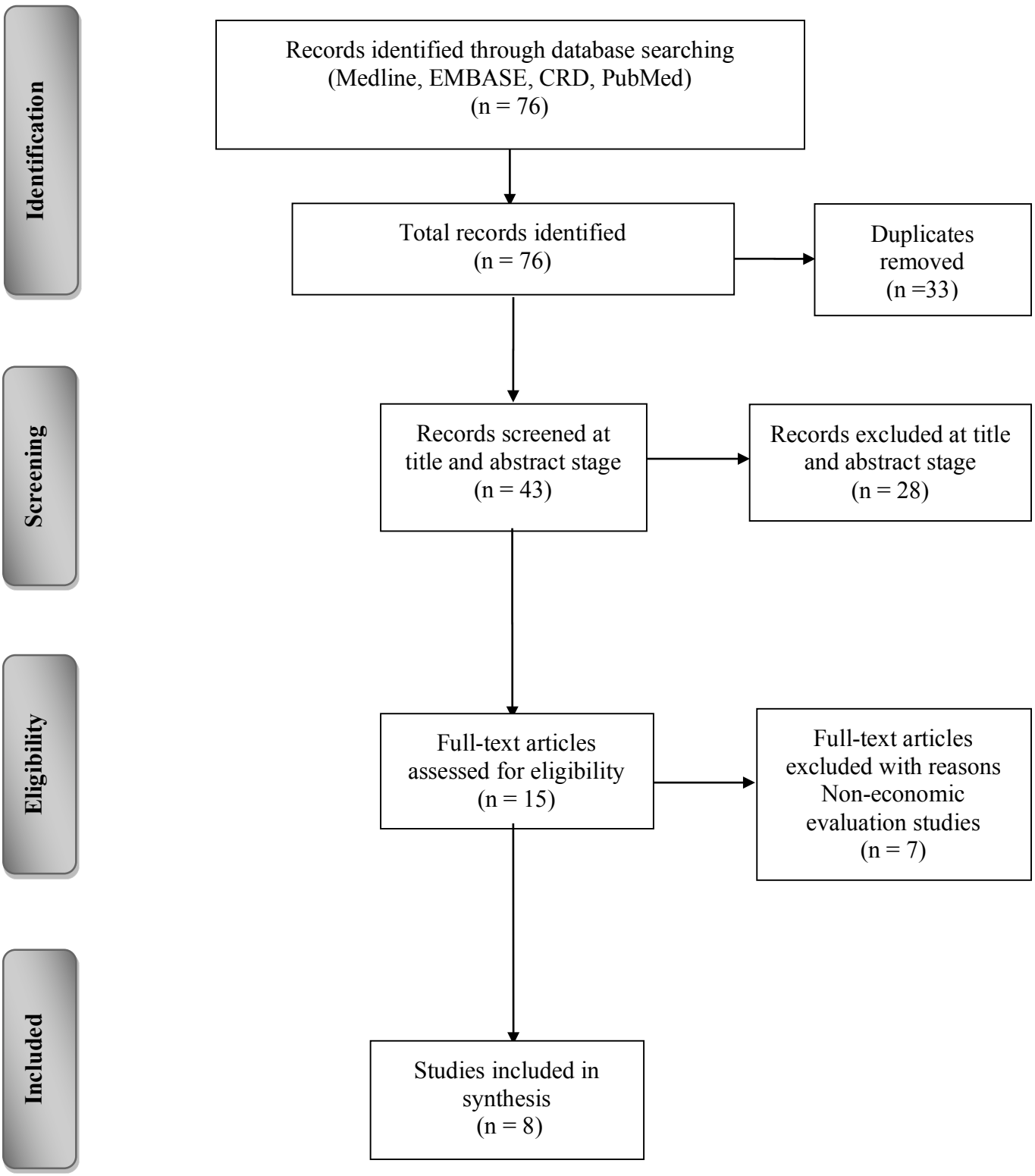
conclusions drawn from these studies. Table 5-1 below shows the quality assessment tool employed in this study.

Table 5-1 Modified Philips checklist criteria for quality assessment of included studies
Is there a clear statement of the decision problem?
Is the objective of the model specified and consistent with the stated decision problem?
Is the primary decision maker specified?
Is the perspective of the model stated clearly?
Are the model inputs consistent with the stated perspective?
Is the structure of the model consistent with a coherent theory of the health condition under evaluation?
Are the sources of the data used to develop the structure of the model specified?
Are the structural assumptions reasonable given the overall objective, perspective and scope of the model?
Is there a clear definition of the options under evaluation?
Have all feasible and practical options been evaluated?
Is there justification for the exclusion of feasible options?
Is the chosen model type appropriate given the decision problem and specified casual relationships within the model?
Is the time horizon of the model sufficient to reflect all important differences between the options?
Do the disease states (state transition model) or the pathways (decision tree model) reflect the underlying biological process of the disease in question and the impact of interventions?
Is the cycle length defined and justified in terms of the natural history of disease?
Source: Adapted from Philips et al, (2006)

5.4 Results

After the removal of duplicates, 43 unique articles were identified for title and abstract screening. Fifteen titles and abstracts were potentially eligible for inclusion. After full text screening, 8 studies were included in the review. The PRISMA flow diagram (Fig 5-1) below illustrates the results of the paper selection process with reason for exclusion noted.

Fig 5-1 The PRISMA flow diagram summarising the search strategy and the paper selection process



5.4.1 Characteristics of included studies

Seven primary studies (Canh et al., 2006; Cook et al., 2008; Cook et al., 2009; Lauria et al., 2009; Musgrove, 1992; Poulos et al., 2004; Shepard et al., 1995) and a systematic review (Watson and Edmund, 2015) were identified. All primary studies retrieved had been included in the systematic review and no new primary studies were identified post the systematic review. The primary studies retrieved and included in this systematic review were published between 1992 and 2009 and the systematic review was published in 2015. Table 5-2 illustrates the characteristics of the included studies.

Table 5-2 Characteristics of included studies

study	Intervention evaluated	Comparator	Type of economic evaluation	Outcome measure	Data source	Perspective adopted
Canh et al., 2006	Vaccination	No vaccination	Cost-benefit analysis	Willingness-to-pay	Field studies	Private sector
Cook et al., 2008	Vaccination	No vaccination	Cost-utility analysis	Cost/DALY averted	Field studies	Public sector and societal
Cook et al., 2009	Vaccination	No vaccination	Cost-benefit analysis	Cost-of-illness avoided & Willingness-to-pay	Field studies	Societal
Lauria et al., 2009	Vaccination	No vaccination	Cost-effectiveness analysis	Cost/case avoided	Field studies	Public sector
Musgrove, 1992	Vaccination	No vaccination	Cost-benefit analysis & cost-effectiveness analysis	Cost/case avoided & maximum cost of vaccination compatible with a net positive benefit	Expert opinion	Public sector
Poulos et al., 2004	Vaccination	No vaccination	Cost-benefit analysis	Cost-of-illness avoided	Field studies	Public sector and societal
Shepard et al., 1995	Vaccination	No vaccination	Cost-utility analysis	Cost/QALY	Expert opinion	Public sector and societal

5.4.2 Quality assessment of included studies

Quality assessment of the included studies showed that 38% of the checklist items were categorized as “yes”, 27% as “no” and 35% as “unclear”. No item was categorized under “not applicable”. The decision problem was clearly stated in all studies and the objectives of the model were also specified which were consistent with the stated decision problem. However, in all the studies, it was unclear if the economic analysis approach used was consistent with the intervention under evaluation. This is because even though the focus of the evaluation in all the studies was the cost-effectiveness of typhoid vaccination, none of the economic evaluation was based on transmission dynamic modelling. None of the studies gave a clear definition of the options under investigation, thus, it was difficult to ascertain whether all feasible and practical options were evaluated. Table 5-3 shows the details of the quality assessment.

Table 5-3 Quality assessment of included studies

Modified Philips checklist criteria	Studies included in the review						
	Canh et al., 2006	**Cook et al., 2008**	**Cook et al., 2009**	**Lauria et al., 2009**	**Musgrove, 1992**	**Poulos et al., 2004**	**Shepard et al., 1995**
Is there a clear statement of the decision problem?	**yes**	**yes**	**yes**	**yes**	**yes**	**yes**	**yes**
Is the objective of the model specified and consistent with the stated decision problem?	**yes**	**yes**	**yes**	**yes**	**yes**	**yes**	**yes**
Is the primary decision maker specified?	**unclear**	**yes**	**unclear**	**yes**	**yes**	**unclear**	**unclear**
Is the perspective of the model stated clearly?	**yes**	**yes**	**yes**	**yes**	**yes**	**yes**	**yes**
Are the model inputs consistent with the stated perspective?	**yes**	**yes**	**unclear**	**yes**	**unclear**	**yes**	**yes**
Is the structure of the model consistent with a coherent theory of the health condition under evaluation?	**unclear**	**unclear**	**unclear**	**unclear**	**unclear**	**unclear**	**unclear**
Are the sources of the data used to develop the structure of the model specified?	**yes**	**yes**	**unclear**	**yes**	**no**	**no**	**unclear**
Are the structural assumptions reasonable given the overall objective, perspective and scope of the model?	**unclear**	**unclear**	**unclear**	**unclear**	**unclear**	**unclear**	**unclear**
Is there a clear definition of the options under evaluation?	**yes**	**yes**	**yes**	**yes**	**no**	**yes**	**yes**
Have all feasible and practical options been evaluated?	**no**	**no**	**no**	**no**	**no**	**no**	**no**
Is there justification for the exclusion of feasible options?	**no**	**no**	**no**	**no**	**no**	**no**	**no**
Is the chosen model type appropriate given the decision problem and specified casual relationships within the model?	**no**	**unclear**	**unclear**	**no**	**unclear**	**no**	**no**
Is the time horizon of the model sufficient to reflect all important differences between the options?	**unclear**	**unclear**	**unclear**	**unclear**	**yes**	**unclear**	**yes**
Do the disease states (state transition model) or the pathways (decision tree model) reflect the underlying biological process of the disease in question and the impact of interventions?	**unclear**	**unclear**	**unclear**	**unclear**	**unclear**	**unclear**	**uncler**
Is the cycle length defined and justified in terms of the natural history of disease?	**no**	**no**	**no**	**no**	**no**	**no**	**no**

5.4.3 Summary of study findings

The cost-effectiveness of typhoid vaccination was the focus of all the primary studies included in this review and none of them was model-based. None of the studies considered typhoid test-treat cost-effectiveness. Two studies were based on values derived from expert opinion (Musgrove, 1992; Shepard et al., 1995), and five studies were based on values derived from field studies (Canh et al., 2006; Cook et al., 2008; Cook et al., 2009; Lauria et al., 2009; Poulos et al., 2004). The evaluations that were based on field studies were noted to share common authorship through collaboration with the Disease of Most Impoverished (DOMI) program. In all studies, it was noted that transmission dynamic modelling was not integrated into the economic analyses although the studies sought to evaluate the cost-effectiveness of typhoid vaccination. For example, analysis by Polous et al (2004) involved dividing the cost of running the vaccination programme in the study population by the economic benefits of Vi Vaccine (which was measured as the avoided *ex-ante* cost of illness in the population). Using a contingent valuation approach, Canh et al (2006) found that a vaccination programme in Hue, Vietnam was more likely to be cost-effective if benefits were measured according to households' willingness to pay (WTP): suggesting a potential for private sector provision of typhoid fever vaccines in Hue, Vietnam. Cook et al (2008) estimated the total costs of immunizing the study population in one year and compared these costs with the effects of immunization on disease burden (cases, deaths, DALYs) over 3 years (duration of the vaccine's effectiveness). Lauria et al (2008) used an optimisation model to determine the vaccine user fees for adults and children that would maximize the number of typhoid cases avoided without increasing public spending. This is the only study that was noted to have included indirect protection (i.e., herd immunity), albeit, using hypothetical values for herd immunity rather than estimates from dynamic modelling. In this particular study, it was shown that vaccine cost-effectiveness was impacted by the level of indirect protection. Whilst the importance of herd

immunity was acknowledged in the other studies, it was noted as having been excluded from their analysis. The absence of evidence was cited as the reason for exclusion. Analysis by Poulus et al. (2004) when conducted from the public sector perspective, showed that a vaccination programme targeted at children under 5 years would be cost-saving. Conducting the same analysis from the societal perspective showed that there were net benefits in other age groups if vaccine cost was moderate and vaccination was carried out in a high incidence setting. Two studies showed that although vaccination with Vi-polysaccharide in both adults and children was unlikely to be cost-effective in a general population setting from the public sector perspective, it would very likely be cost-effective in a high incidence setting (Musgrove, 1992; Shepard et al., 1995). In these studies, it was established through sensitivity analysis that the main drivers of cost-effectiveness were vaccine cost, vaccine duration of protection, case fatality rate and vaccine effectiveness. Indirect protection was not assumed in these studies: therefore, the effect of herd immunity on cost-effectiveness could not be appraised. It is worth stating that whilst vaccination programmes were predicted to reduce typhoid incidence, it is not easy to draw this conclusion without the studies having included some measure of herd-immunity in the analysis. It was noted that short or medium-term vaccination programs are unlikely to be effective in the elimination of the disease without measures aimed at reducing the ongoing force of infection (the rate at which susceptible individuals become infected) such as asymptomatic carriers.

5.5 Discussion

This review examined previous typhoid economic evaluations, with particular focus on how modelling for typhoid interventions had been approached. The review showed that there have been relatively few economic evaluations that have focussed on typhoid fever, all of which have focused on the cost-effectiveness of typhoid vaccination although none of them was

model-based. There are a number of reasons why a new systematic review may be considered. However, before undertaking a systematic review it is necessary to check whether there are already existing or ongoing reviews, and whether a new review is justified. Although a systematic review on the topic area was identified through the paper selection process, the focus of that systematic review was to examine typhoid vaccine economic evaluations for their potential contributions to inform disease control. This differs markedly from the focus of the systematic review in this thesis, thus the justification for undertaking this systematic review.

Vaccination acts by stimulating the immune system of the host and operates directly by reducing the number of susceptible individuals in the population and indirectly via "herd immunity". When evaluating typhoid vaccine effectiveness, it is key that both these effects are captured. Transmission dynamic models are well suited for capturing both effects: thus, they are appropriate when evaluating the cost-effectiveness of typhoid vaccination (Keeling and Rohani, 2011). However, as noted, this was not the case in the studies included in this analysis, consequently, indirect protection was omitted in the analysis. Therefore, it was not possible to appraise the indirect effect of vaccination in their analysis. The implication is that this may lead to an underestimation or overestimation of the true benefits of vaccination and may result in inappropriate decision-making. Hence, transmission dynamic modelling should be integrated into cost-effectiveness analysis when estimating the true value of typhoid vaccines for these evaluations to be useful to policy making. However, under certain circumstances, using a static model based on only direct protection may be a reasonable approximation. For example, this will be true if the new cases that are prevented as a result of vaccination make a relatively little contribution to the rate at which susceptible individuals acquire the disease (Watson and Edmund, 2015).

None of the studies included in this review focussed on typhoid test-treat strategies. Thus, the review shows that the impact of typhoid test-treat strategies has not been explored using modelling. A notable feature of typhoid fever is chronic carriers (1-5%) who continue to shed the organisms via either their stool or urine. Thus, sustaining the occurrence of the disease in endemic settings (WHO, 2003). Furthermore, short-term convalescent patients may also contribute to disease transmission through faecal shedding in an endemic setting. Early accurate diagnosis and treatment of a case of typhoid fever (new case, short-term convalescent, or chronic carrier) focused on curtailing shedding can potentially affect onward transmission. However, the potential benefits of treatment following early accurate diagnosis in the prevention of onward transmission of the infection in an endemic setting has been little studied. Thus, there is no evidence to inform the extent to which treatment affects transmission. Therefore, when evaluating the the cost-effectiveness of an intervention for typhoid fever (such as test-treat strategies) where the emphasis is on improving health outcomes (as is the case in this work) rather than benefits to the population as a result of treatment impacting on transmission, then a static model may suffice (Drake et al., 2016). Indeed, the role of static models to evaluate the cost-effectiveness of rapid diagnostic testing in resource-limited settings have been shown in other disease areas where the focus was to improve direct health outcomes without necessarily impacting disease transmission (Tawiah et al., 2016; Hansen et al., 2017). One parameter that has been found to be a major driver of cost-effectiveness in the field of typhoid economic evaluations is "incidence". Most studies have focussed on the estimation of incidence thresholds to guide policy decision-making. However, it has also been shown that covariates such as case fatality rate, antimicrobial resistance, and access to quality health care are critical but uncertain parameters that drive the incidence threshold (Lo et al., 2018). Whilst static models have their limitations, they can certainly be used to assess the importance of these parameters on driving the conclusions from models such as these.

5.6 Conclusion

Under certain circumstances, a static model or a dynamic model may be appropriate in the evaluation of an intervention for typhoid fever. Typhoid test-treat modelling represents a grey area where further work is needed. The next chapter (Chapter 6) builds on this by demonstrating the use of static models to evaluate the potential cost-effectiveness of rapid diagnostic testing for typhoid fever in Ghana.

Chapter 6: An early cost-effectiveness study of a hypothetical rapid test for typhoid fever diagnosis in Ghana

This chapter brings together and employs all the knowledge and findings from the previous chapters in an early cost-effectiveness analysis.

6.1 Background

In Ghana, there are some issues of concern with the diagnosis of typhoid fever (see Chapter 1, p. 1-2). A quick turnaround time (leading to early diagnosis and treatment), better accuracy and availability are some of the potential benefits that a rapid diagnostic test could offer over current testing strategies in Ghana (see Chapter 4, p. 82). Therefore, the development of a test for typhoid fever in the Ghanaian setting which has better accuracy than the current available tests, and which is simple and affordable with a quick turnaround time to enable prompt implementation of an appropriate effective treatment has an obvious attraction. The aim of this chapter was to examine the potential cost-effectiveness of the HT-test in Ghana, with specific focus on the estimation of the maximum price and the minimum test performance required for the HT-test to be cost-effective. In this early cost-effectiveness analysis, the HT-test is the index test (the test to be evaluated). The comparator interventions (in this case the reference standards) against which the performance of the HT-test is compared are the culture testing pathway and the Widal testing pathway which were identified in Chapter 4 as being the main diagnostic tests used for typhoid fever in Ghana (see Chapter 4, p. 76).

6.2 Methods

The patient population considered in this analysis is made up of individuals presenting with symptoms suggestive of typhoid fever (diagnosis is unknown at presentation) at a health care facility in Ghana. A time horizon of 180 days was chosen because in this analysis all pathways eventually lead to successful treatment and in such cases the appropriate time horizon needs to be as long as the longest treatment pathway. In this instance, the longest treatment pathway was 34 days. Thus, 180 days was chosen to be an appropriate time over which patients will benefit from the effects of testing and treatment for typhoid fever (Lo et al., 2018). Also, by taking the time horizon to 180 days, it would allow other feasible options to be added to the model which could have a longer treatment time. The Ghanaian national health service perspective was adopted for the analysis because the health service is the direct payer for health care services and the aim is to support decision making for the health service (EUnetHTA, 2015). No discounting of cost and benefits was undertaken because the time horizon did not exceed 12 months (Drummond et al., 2005). The findings are expressed in terms of cost, quality-adjusted life years (QALYs) and net-monetary benefit (NMB). Here, the NMB is defined for each intervention as:

NMB = QALYs gained * willingness to pay (WTP) for a QALY – cost of intervention.

In Ghana, a cost-effectiveness threshold of less than $104-$951/QALY is generally considered to be cost-effective by policy makers (Woods et al., 2016). For this analysis, the upper limit of the cost-effectiveness threshold value ($951/QALY) was chosen because the lower limit is practically too low.

6.2.1 Model Choice

Evaluating the potential value of the HT-test (or indeed any new technology at its early stage of development) requires the use of models because of the lack of data describing the effectiveness of an intervention during its early stages of development. An important first step was to decide the type of model to use (static or dynamic); i.e., whether transmission of infection between individuals needed to be captured in the model? Although it is acknowledged that typhoid fever is an infectious disease and transmission dynamic modelling would be better suited for its evaluation, a static model rather than a dynamic model was adopted in this study. Underpinning this choice is the fact that, unlike vaccination where its role in reducing the incidence of typhoid (directly and indirectly which is best captured by a dynamic model) has been demonstrated (Watson and Edmunds, 2015), there is no evidence to inform the extent to which early diagnosis and treatment affects onward transmission as this has been little studied (see Chapter 5, p. 101). In the case of this analysis, the emphasis was on evaluation of the direct benefits of testing and treatment to an individual (i.e., survival and quality of life) rather than benefits to the population via prevention of onward transmission. Furthermore, all the patients in this model are assumed to be treated successfully, thus, they will not contribute to onward transmission although there is the issue with a few days' delay (3 days) before they are cured. However, because the means of transmission is not direct person-to-person (Akullian et al., 2015), even a few days' delay will not be likely to make much difference to the overall force of infection. Also, when interpreting results, the fact that including transmission effects will only increase the health gains from treatment, any intervention shown to be cost-effective is even more likely to be cost-effective in real life. Thus, under these circumstances, a static model was considered appropriate for the analysis.

The type of static model to employ (i.e., decision tree or Markov model) depends on the nature of the clinical condition being evaluated (i.e., acute or chronic). Decision trees are used for "one off" decisions and are particularly suited to acute clinical conditions, once-only diseases, and short-term diagnostic or screening decisions whereas Markov models are more suited for modelling the progression of chronic disease (Drummond et al., 2005). As discussed in Chapter 2, typhoid fever is an acute condition. Therefore, in this study, decision trees were constructed to undertake the evaluation.

6.2.2 Structuring the model

The ensuing sections outline the assumptions and the steps in constructing and analysing the decision trees used in this analysis.

6.2.2.1 Position of HT-test on the clinical pathway

In this analysis, the value of the hypothetical test in each of the two care pathways was considered separately due to differences in existing pathways. A key determination in test evaluation is the best placement and role of a new test on existing pathways. Therefore, as a comparator for the two current tests (i.e., the Widal care pathway and the culture care pathway used in Ghana), the HT-test was evaluated in three possible roles (replacement, triage or add-on), resulting in the construction of six decision trees. In this analysis, the emphasis of the evaluation was to determine in each role the maximum cost of the HT-test at a given sensitivity and specificity pair (using threshold analysis) at which it was still cost-effective at a WTP threshold of $951/QALY. There was no intention to compare the different roles with each other in the same pathway. The index test (the HT-test) can only appear in the comparisons once. Thus, there could be multiple current practice options but only one hypothetical testing strategy which explains why six different trees were constructed and not two (one of the two current

tests and one of the three possible roles of the new tests in each one). This was done to provide the policy maker (Government of Ghana) with a complete analysis of the potential benefits of the HT-test in each of the two pathways and in each of the different roles to support informed decision-making. Each model has two arms (current practice and the hypothetical testing strategy). The current practice arm is the same in all three scenarios in both pathways, but the hypothetical testing strategy arm differs depending on the role of the test on the test-treat pathway (replacement, triage or add-on). The test-treat pathway for a suspected typhoid case is as shown in the different decision trees. In the replacement role, the HT-test is a potential substitute for current tests. In the triage role, the HT-test is used to decide who should receive further testing with current tests, while in the add-on role the HT-test provides additional diagnostic information following the current test. Fig 6-1 shows the model for the HT-test vs. Widal test in the replacement Role. The other five models are shown in Appendix 5. Fig 6-2 shows the Model structure showing the comparison between HT-test vs. Widal test in the different roles.

Fig 6-1 Model structure for evaluating the HT-test vs. Widal test in the replacement role

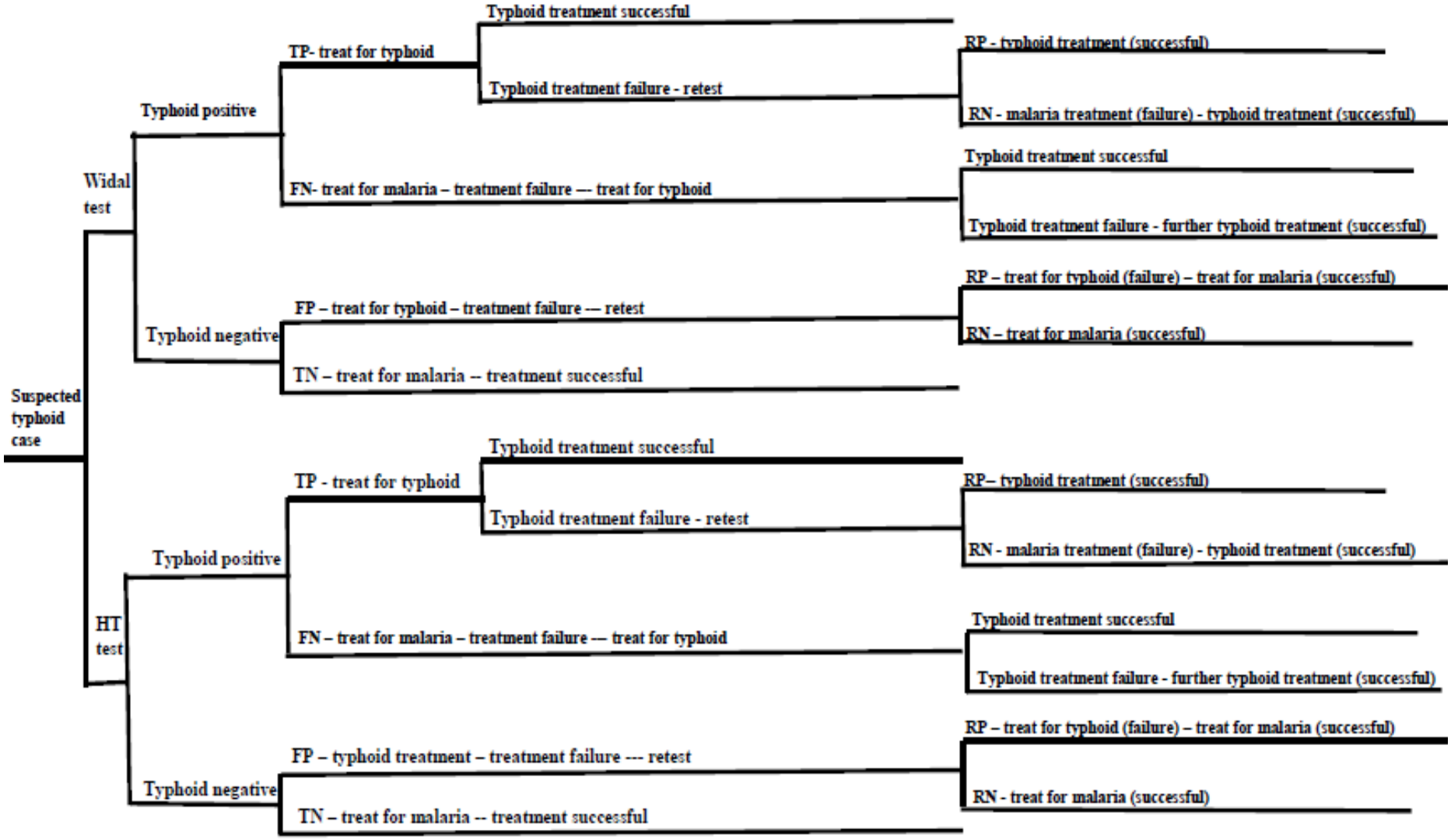


Fig 6-2 Model structure showing the comparison between HT-test vs. Widal test in the different roles

6.2.2.2 Test accuracy

In this analysis, each test is considered to provide either a positive or negative test result. This allows the application of sensitivity and specificity values to be considered in the analysis. The resulting possible test statuses for patients following the different pathways are shown in Table 6-1 below.

Table 6-1 Description of the application of test statuses used in the models

Status	Description
TP	True Positive (has typhoid and tests positive for it)
TN	True Negative (does not have typhoid and tests negative for it)
FP	False Positive (does not have typhoid but tests positive for it)
FN	False Negative (has typhoid but tests negative for it)

Given that the HT-test is a hypothetical test, there was no information on the sensitivity and specificity of the test. For the Widal test and blood culture, following a detailed search of the literature, a meta-analysis study of typhoid diagnostic accuracy studies was found and used to inform their test accuracy parameters as shown in Table 6-2 below.

Table 6-2 Sensitivity and specificity values for Widal test and blood culture

Parameter	Estimate	Data source
Sensitivity of Blood culture (CI)	0.68 (0.52 - 0.81)	Storey et al., (2015)
Specificity of Blood culture (CI)	0.75 (0.35 – 0.94)	Storey et al., (2015)
Sensitivity of Widal test (CI)	0.69 (0.61- 0.75)	Storey et al., (2015)
Specificity of Widal test (CI)	0.83 (0.77- 0.88)	Storey et al., (2015)

6.2.2.3 Model assumptions

In order to implement the models, the following assumptions were made:

- For a patient presenting with suspected typhoid, if the patient does not have typhoid, then it is assumed they have malaria. Underpinning this assumption is the fact that, the presenting symptoms for both conditions are similar (Wain et al., 2015) and the most common differential diagnosis for typhoid fever is malaria (see Chapter 4, p. 81). This is the fundamental assumption of the model.
- Test results obtained from retesting a person who has had treatment failure with the same test, are independent of those from the first test. Underpinning this assumption is the fact that, as part of the typhoid test-treat protocol in Ghana, patients are retested with the same test on their review appointments and further treatment decisions are made based on the test results (see Chapter 4, P. 81). Thus, the implicit assumption that errors are purely random in second test.
- If the first round of typhoid treatment fails in individuals who actually have typhoid, then the second round will always be successful. Underpinning this assumption is the fact that, ciprofloxacin has been found to be highly active against *Salmonellae* in vivo and repeated treatment is usually associated with a 100% cure rate (Alam et al., 1995).
- For any patient that has malaria, malaria treatment is successful. This is justified because artemisinin-based combination treatments (ACTs) (the first line drug for uncomplicated malaria in Ghana) are rapid and effective (effectiveness usually exceeds 95%) (Nosten and White, 2007).
- There is no uncertainty about the length of treatment for typhoid and malaria (14 days and 3 days respectively). These are informed by the standard treatment guidelines in Ghana (GNDP, 2014).

- All patients experiencing treatment failure return for retesting immediately after completion of a treatment course. Underpinning this assumption is the fact that as part of the clinical practice in Ghana, patients are given review dates which falls on the last day of a treatment course and are highly encouraged to return for reassessment and this seems to be the case (see Chapter 4, page 81). Thus, the implicit assumption made.
- Delays on the test-treat pathway (i.e., delay receiving test results, delay receiving treatment) are negligible. Underpinning this assumption is the fact that there are available tests and treatments for typhoid fever in Ghana (GNDP, 2014), thus, once patients have reported to the healthcare facility they are tested and an appropriate treatment is given (see Chapter 4, p 81).
- Patients experience the same utility at the different phases of treatment but return to perfect health once treatment has been completed. This assumption was made because of the paucity of evidence in the literature on utility estimates to inform the different phases of treatment.
- Utility when experiencing typhoid symptoms is the same as when experiencing malaria symptoms as both diseases are similar in the acute phase of illness (Mutua et al., 2015).

6.2.2.4 Individuals tested who have typhoid

Any action taken is dependent on the information available to the clinician at that point in time (i.e. test results and response to treatment). For example, as shown in the current practice arm (Fig 6-1), individuals presenting with symptoms suggestive of typhoid who test positive and actually have typhoid (True positive - TP) are given a first round of typhoid treatment, some of whom will be successfully treated and others not. TPs experiencing treatment failure are retested and re-test TPs are given a second round of typhoid treatment which in the model is assumed to be successful. Retested negatives (re-test FN) will be given malaria treatment on the basis that typhoid treatment has been unsuccessful. However, because re-test FN patients actually have typhoid, the malaria treatment will be ineffective. They are then given a second round of typhoid treatment following the malaria treatment and this second round of treatment is assumed to be successful.

6.2.2.5 Individuals tested who do not have typhoid

As shown in the current practice arm (Fig 6-1), individuals presenting with symptoms suggestive of typhoid who test negative and do not have typhoid (and assumed to have malaria) (True negative - TN) are given treatment for malaria which in the model is assumed to be successful. Individuals presenting with symptoms suggestive of typhoid who test positive but do not have typhoid (False positive - FP) are treated for typhoid based on their test results. However, because they do not have typhoid, typhoid treatment will be ineffective and the patients are retested. Test positive patients as a result of a re-test (re-test FP) are given a second round of typhoid treatment which will be ineffective for the same reason. Based on two typhoid treatment failures, they are given treatment for malaria which is assumed to be successful. Test negative patients as a result of a re-test (re-test TN) are given treatment for malaria which is assumed to be successful in the model.

6.2.3 Parameterization

A range of secondary sources was utilized to parameterize the model. These secondary sources were identified through extensive searching of published literature, websites of government departments in Ghana, and medical records in Ghana. Table 6-3 shows the probabilities used in the economic analysis along with the data source.

Table 6-3 Probability estimates utilized in the economic analysis

Parameter	Estimate	Data source
Probability of successful first typhoid treatment	0.95	Chandey and Multani, (2012)
Prevalence of typhoid in the patient population presenting	0.48	Afoakwah et al., (2011)

Given that this is an early economic evaluation, there was a lack of data to inform estimates of the parameters describing the characteristics of the HT-test, thus these parameters were subject to extensive sensitivity analysis to draw conclusions about the characteristics of this test necessary for it to be cost-effective.

6.2.3.1 Measurement of effects

The primary outcome measure was the QALY. Total QALYs were estimated for each branch based on the parameters described in Table 6-4 below.

Table 6-4 Parameter estimates used to inform QALY values		
Item	**Estimate**	**Source**
Utility when experiencing typhoid symptoms (CI)	0.867 (0.81-0.912)	Lo et al., (2018)
Perfect health utility	1	By definition: Drummond et al., (2005)
Time horizon of model (days)	180	Lo et al., (2018)
Number of days per typhoid treatment	14	GNDP, (2014)
Mean recovery time for successful typhoid treatment (SD)	3.68 (0.92)	Chandey and Multani, (2012)
Number of days per malaria treatment	3	GNDP, (2014)
Mean recovery time for successful malaria treatment (SD)	1.83 (0.95)	Grynberg et al., (2015)

For every round of unsuccessful treatment, the total number of days per treatment is included in the estimation, and for every successful treatment, the mean recovery time for successful treatment is also included. For example, total QALYs for a branch where a patient receives two rounds of typhoid treatment (i.e., one unsuccessful and one successful) was calculated as, $(U_1T_1) + (U_1T_2) + (U_2T_3)$ / 365 days.

Where U_1 is the utility when experiencing typhoid symptoms.

T_1 is number of days per typhoid treatment.

T_2 is mean recovery time for successful typhoid treatment.

U_2 is perfect health utility.

T_3 is the number of days in perfect health after two typhoid treatments.

The time horizon (TH) has been incorporated into the above calculation as follows:

$T_3 = TH - (T_1+T_2)$

6.2.3.2 Measurements of cost

Total cost for each branch was estimated based on the perspective of the Ghanaian national health service and for the 2018 cost year. All costs were converted to US dollars by using the exchange rate (Gh₵1 = $0.225) adopted from the website of the Bank of Ghana at the time of analysis (29/03/2018) (BoG, 2018). Table 6-5 shows the costs used in the economic analysis.

Table 6-5 Cost parameters and references

Item	Unit cost (Gh₵)	Unit cost ($)	Data source
Widal test	5.67	1.28	NHIS tariffs, (2016)
Blood culture	10.18	2.29	NHIS tariffs, (2016)
Ciprofloxacin	6	1.35	NHIS tariffs, (2016)
Artemether/Lumefantrine	3.5	0.79	NHIS tariffs, (2016)
OPD tariff per visit	16.07	3.62	NHIS tariffs, (2016)

Ciprofloxacin and Artemether/Lumefantrine are the first line drugs of choice for typhoid and malaria treatment respectively in Ghana. The total cost for a branch where a patient has visited the health facility once, received the Widal test and then typhoid treatment is estimated as follows:

cost of Widal test + cost of ciprofloxacin + OPD tariff per visit = $ 6.25

The OPD tariff is an amount reimbursable to the service provider per visit to cover other costs like staff time and consultation. On any pathway where initial treatment is unsuccessful, there will be more than one visit and, in such instances, the OPD tariff is multiplied by the number of visits and added to the other costs incurred in order to estimate the total cost for that branch.

6.2.4 Analysing the decision tree

After the model had been populated, the decision tree was averaged out and "rolled-back" to obtain the expected costs and QALYs gained for each strategy. These are the sum of the products of the estimates of the probability of events and their outcomes.

6.2.5 Dealing with uncertainty

Sensitivity analyses were conducted to test the robustness of parameter estimates, determine the range of parameters which have the greatest impact on cost-effectiveness, and determine how sensitive the results are to changes in model structure. Given that this is an early economic evaluation of a hypothetical test, there was no information on test accuracy or cost of the new test. Thus, each of these parameters was a separate focus of sensitivity analysis. Headroom analysis (i.e., maximum cost at which the new technology is still cost-effective at a given willingness-to-pay threshold) using NMB was conducted. The price at which the NMB of using the HT-test equals the NMB of using each of the comparator tests gives information on the headroom price of the HT-test compared to each of the comparator tests. Furthermore, sensitivity and specificity pairs for the HT-test were varied and the headroom price for each pair determined. This analysis gave information on how good the test needs to be (in terms of accuracy), and at what price, for it to be cost-effective. By varying accuracy pairs for the HT-test, the incremental effectiveness for each accuracy pair could be determined and information on which accuracy pairs provide the most benefit to patients within cost-effectiveness constraints could be identified.

Probabilistic sensitivity analysis (PSA) was also conducted. A probability distribution was defined for each uncertain parameter (i.e., sensitivity and specificity of Widal test, sensitivity

and specificity of culture, prevalence of typhoid fever in the patient population presenting, probability of successful first typhoid treatment, mean recovery time for successful typhoid treatment and mean recovery time for successful malaria treatment). A beta distribution was assigned to all uncertain parameters except the test accuracy parameters of the comparator tests. Table 6-6 shows the beta distributions assigned to each of the uncertain model parameters.

Table 6-6 Parameters of beta distribution

Model parameters with uncertainty	Parameters of beta distribution	
	a	b
Mean recovery time for successful typhoid treatment (This is a fraction of the length of treatment for typhoid)	9.95	28.32
Mean recovery time for successful malaria treatment (This is a fraction of the length of treatment for malaria)	0.85	0.55
Prevalence of typhoid in the patient population presenting	22	24
Probability of successful first typhoid treatment	38	2
Utility when experiencing typhoid symptoms	145.48	22.32

Sensitivity and specificity are correlated and therefore required a different method that accounts for the correlation between these parameters (Harbord and Whiting, 2009). Data from a meta-analysis of typhoid diagnostic accuracy studies (Storey et al., 2015) was utilized to obtain the hierarchical summary receiver operating characteristic (HSROC) curves for Widal test and blood culture. Using the Stata function metandi to obtain the HSROC curve and the following equation that links sensitivity to any given specificity, it was possible to take a value of specificity and calculate a corresponding sensitivity along the HSROC curve.

$$logit(sensitivity) = \Lambda e - \beta/2 \; - e - \beta \; logit(specificity)$$

Where lambda (Λ) is the mean of the accuracy parameter.

Beta (β) is the shape parameter.

For the PSA, lambda and beta were described by a normal distribution using the mean and standard error described in the Stata output. These distributions were then sampled for the PSA. The model was run 1000 times, each time randomly selecting a value for each parameter from their respective distributions. Mean costs and mean QALYs were calculated by averaging across all 1000 simulations. Cost-effectiveness acceptability curves (CEACs) were drawn to gain a greater understanding of realistic price and test accuracy parameters.

6.2.6 Value of information (VOI) analysis

VOI analysis was conducted to provide insights into the value of conducting further research from the perspective of the health care provider in Ghana. In the absence of information on test accuracy parameters of the HT-test and specific guidance on VOI analysis at this early stage of analysis, the assumption was made that the HT-test has the same accuracy characteristics as the Widal test (the better of the two comparator tests in terms of accuracy). This approach was taken in order to reflect in the results some sense of the value of resolving this uncertainty as the accuracy of the HT-test is unknown. Thus, the accuracy parameters for the Widal test (i.e., sensitivity of 0.69 and specificity of 0.83 with their associated uncertainty) were used to inform the test accuracy of the HT-test in both comparisons (i.e., the HT-test vs. Widal test and the HT-test vs. blood culture). The VOI analysis was repeated for different prices in each role as informed by the results of the headroom analysis. The assumed sensitivity and specificity of the HT-test were incorporated into the PSA to capture the variations around the HT-test accuracy characteristics in the analysis. To do this, the distributions for the Widal test were

sampled 2,000 times. The first 1,000 samples were used to describe the uncertainty around the Widal test characteristics (sensitivity and specificity) and the second set of 1,000 samples for uncertainty around the HT-test characteristics. When thinking about the overall value of removing decision uncertainty, one needs to consider the number of people affected by the decision annually and how long the current comparison will remain relevant. Kim et al (2017) reports that the annual incidence of typhoid in Ghana in 2016 was 60,892, thus, in this analysis, the number of people considered to be affected by the decision annually was 60,000. Also, it is assumed that the HT-test is likely to be relevant for five years before new technologies become available, thus, a decision relevant time horizon of 5 years was adopted in this analysis. In the absence of an accepted discounting rate in Ghana, a discounting rate of 3.5% (as applied in the UK) was adopted in this analysis to estimate the present value of the overall value of removing decision uncertainty over the specified decision relevance time horizon of 5 years. The formula below was used to estimate the total discounted incidence over 5 years. This value was multiplied by the expected value of perfect information (EVPI) per person in each scenario to obtain the overall value of removing decision uncertainty over 5 years.

$N * (1+ 1/(1+r) + 1/(1+r)^2 + 1/(1+r)^3 + 1/(1+r)^4)$

Where N is the annual incidence (in the case of this analysis N= 60,000)

r is the discounting rate (in the case of this analysis r = 0.035)

The total discounted incidence used in this analysis is 280,385.

In VOI analysis, one may also want information on the parameters causing most of the decision uncertainty and the potential value of reducing the uncertainty by collecting more data. This is known as expected value of partial perfect information (EVPPI). In this study, EVPPI was estimated using the Sheffield Accelerated Value of Information (SAVI) software (Strong et al., 2014). Information about the model (i.e., model name, WTP threshold, number of patients affected by the decision each year, decision relevance time horizon, units used for cost and

benefits and name of jurisdiction/country) were inputted into the software. Then, the PSA samples of parameters, costs and effects were imported into the software as three csv files and the software was run to generate a report on the EVPPI. The potential limitation of using the SAVI software is that, it does not incorporate a discounting rate and therefore, the estimated EVPPI over the relevance time horizon is a direct multiplication of the per person EVPPI by the time horizon. This is inappropriate and a discounting rate needs to be introduced into this calculation in order to estimate the actual EVPPI value over the relevant time horizon. This particular limitation was not considered applicable in this analysis because, it was identified from the outset and the overall EVPPI for the relevant time horizon was calculated by multiplying the per person EVPPI with the estimated total discounted incidence over 5 years (i.e., 280,385).

6.3 Results

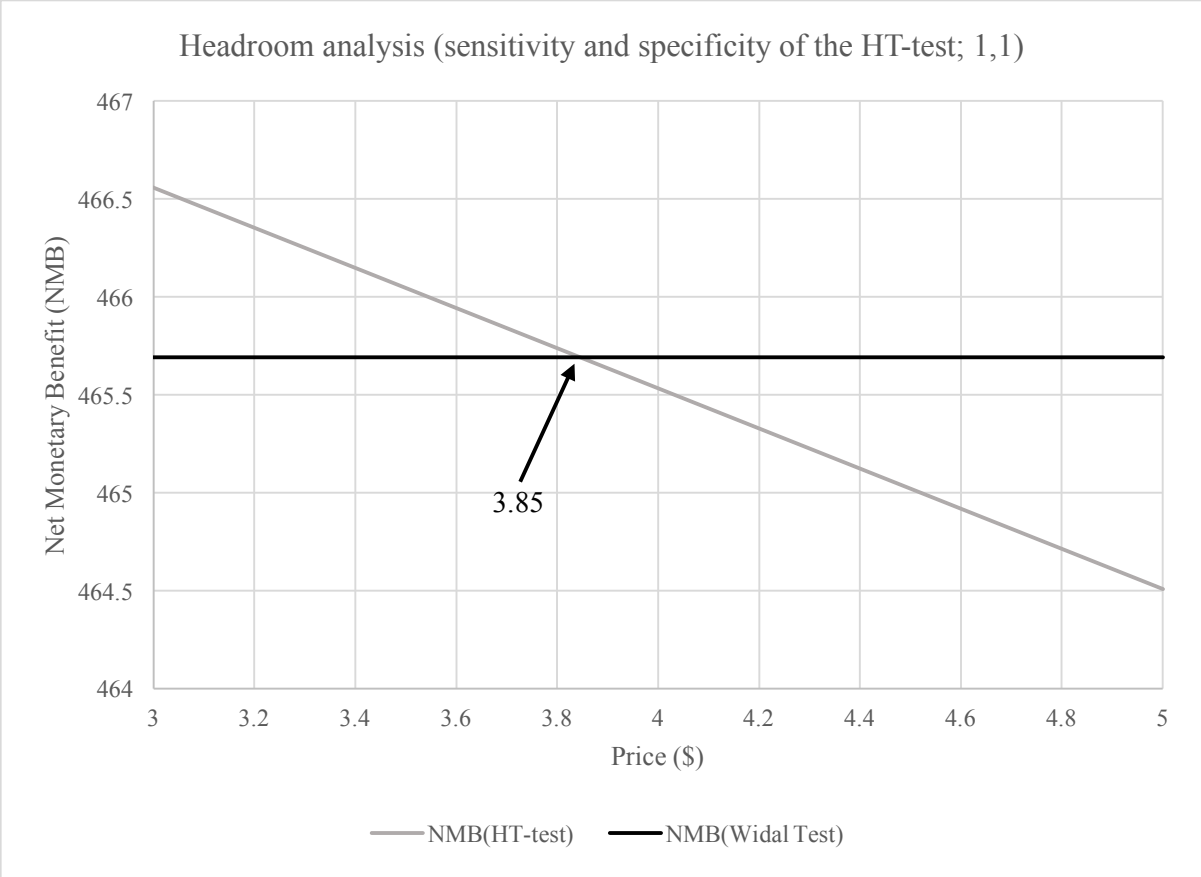
The results are presented in two parts. The results of HT-test vs. Widal test are presented first, followed by the results of the HT-test vs. blood culture. In each case, the results are presented in the following order. First, graphs of NMB of the HT-test and each of the comparator tests (which gives information on headroom price) are presented. These are followed by tables of incremental effectiveness and the maximum price at which the HT-test is still cost-effective (at a threshold of $951/QALY) at each sensitivity and specificity pair for the HT-test against each of the comparator tests. Then CEACs and a table summarising the results for the comparisons between the HT-test and each of the comparator tests is presented. Finally, VOI analysis results are given.

6.3.1 Comparisons of the HT-test vs. Widal test

6.3.1.1 Headroom analysis using NMB (the HT-test vs. Widal test)

Taking all parameters at baseline and using NMB for each intervention at a WTP for a QALY of $951, the NMB for the HT-test and Widal test with variation in the price for the HT-test in the replacement role is shown in Fig 6-3. As shown in Fig 6-3, at a WTP of $951 for a QALY, the HT-test can be priced up to $3.85 and still be cost-effective against the Widal test in the replacement role. The maximum prices at which the HT-test can be priced and still be cost-effective against the Widal test in the triage and add-on roles are $1.94 and $1.47 respectively. The figures showing NMB for triage and add-on roles are presented in Appendix 6.

Fig 6-3 Headroom of the HT-test vs. Widal test at a WTP of $951/QALY (Replacement role)



6.3.1.2 Incremental effectiveness (QALYs) varying sensitivity and specificity pairs for the HT-test vs. Widal test

The following set of tables show the incremental effectiveness and the maximum price at which the HT-test is still cost-effective (at $951/QALY) at each sensitivity and specificity pair for the HT-test vs. Widal test in the replacement role. The tables for the triage and add-on roles are presented in Appendix 7. When examined together, these tables provide valuable and useful information critical to the understanding of how accurate the HT-test needs to be, and at what price, for it to be cost-effective.

NV in the tables represent negative values for price and implies that the HT-test cannot be cost-effective at these sensitivity and specificity pairs and the end user of the test would have to be paid for using the test. Shaded black cells represent cases where either the test is completely useless, or it would be better to take the reverse of the test results.

Table 6-7 Incremental effectiveness (QALYs) at each sensitivity and specificity pair for the HT-test vs. Widal test (Replacement role)

Sensitivity \ Specificity	100	90	80	70	60	50	40	30	20	10	0
100	0.0010	0.0007	0.0004	0	-0.0005	-0.001	-0.0016	-0.0022	-0.0029	-0.0036	
90	0.0010	0.0007	0.0003	-0.0001	-0.0005	-0.0011	-0.0016	-0.0022	-0.0029		
80	0.0009	0.0006	0.0003	-0.0001	-0.0006	-0.0011	-0.0017	-0.0023			
70	0.0009	0.0006	0.0002	-0.0002	-0.0007	-0.0012	-0.0017				
60	0.0008	0.0005	0.0002	-0.0003	-0.0007	-0.0012					
50	0.0007	0.0004	0.0001	-0.0003	-0.0008						
40	0.0007	0.0004	0	-0.0004							
30	0.0006	0.0003	0								
20	0.0006	0.0003									
10	0.0005										
0											

Table 6-8 Maximum price ($) at which the HT-test is still cost-effective at each sensitivity and specificity pair for the HT-test vs. Widal test (Replacement role)

Sensitivity \ Specificity	100	90	80	70	60	50	40	30	20	10	0
100	**3.85**	**3.13**	**2.40**	**1.64**	*0.86*	*0.06*	NV	NV	NV	NV	
90	**3.59**	**2.89**	**2.16**	*1.41*	*0.64*	NV	NV	NV	NV		
80	**3.33**	**2.64**	**1.92**	*1.18*	*0.42*	NV	NV	NV			
70	**3.07**	**2.39**	**1.69**	*0.95*	*0.2*	NV	NV				
60	**2.82**	**2.15**	**1.45**	*0.73*	NV	NV					
50	**2.56**	**1.91**	**1.22**	*0.5*	NV						
40	**2.31**	**1.67**	**0.99**	*0.28*							
30	**2.06**	**1.43**	**0.76**								
20	**1.81**	**1.19**									
10	**1.56**										
0											

Values in **bold** format represents values of price at sensitivity and specificity pairs at which the HT-test is still cost-effective and has a positive incremental effect, and would normally be of interest to a decision maker.

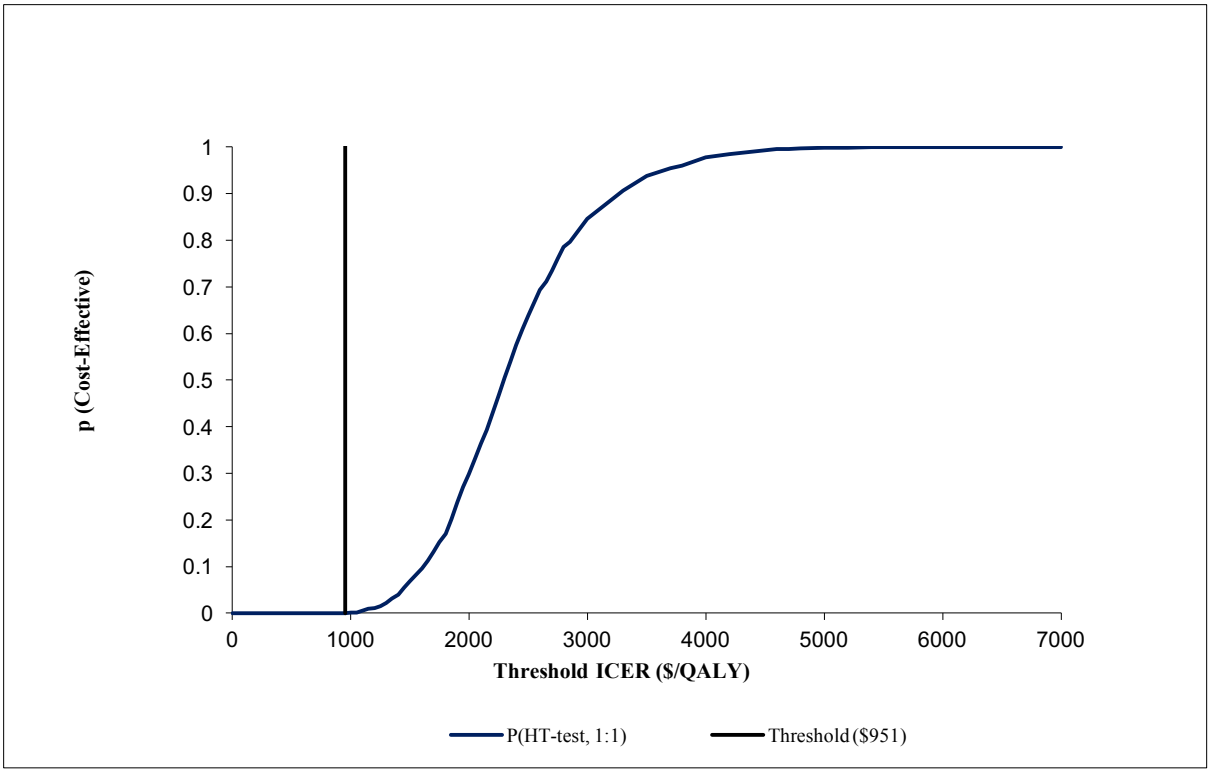
Values in *italic* format represent values of price at sensitivity and specificity pairs at which the HT-test is still cost-effective but is slightly less effective than the Widal test.

As shown in the set of tables (Tables 6-7 & 6-8), the HT-test can be priced up to $3.85 and still be cost effective against the Widal test in the replacement role when the HT-test has both sensitivity and specificity of 100%. In this case, the incremental effectiveness is 0.0010. When the HT-test has both sensitivity and specificity of 70%, it can be priced up to $0.95 and still be cost-effective, however, here the incremental effectiveness is -0.0002. This means that as the accuracy of the HT-test decreases, the maximum price of the HT-test for it to be cost-effective against the Widal test in the replacement role reduces with a decrease in incremental effectiveness. The same observation is made for the triage and add-on roles.

6.3.1.3 Cost-effectiveness acceptability curves for the HT-test vs. Widal test (replacement role)

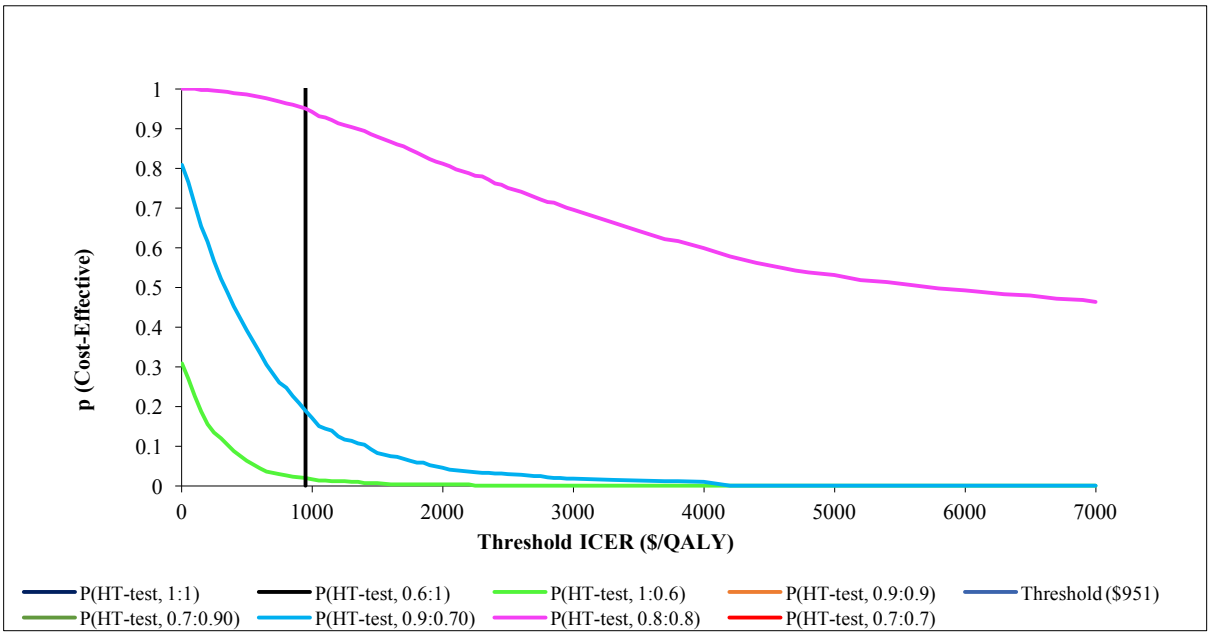
Using a reasonable price range for the HT-test of $1.00 to $4.00 that was informed by findings shown in Table 6-7 and Table 6-8, the CEACs for the PSA of the HT-test vs. Widal test in the replacement role are presented below. The set of figures for the triage and add-on roles are presented in Appendix 8. In the CEACs, the price is fixed whereas sensitivity and specificity are varied using the following pairs of values (1:1; 0.6:1; 1:0.6; 0.9:0.9; 0.7:0.9; 0.9:0.7; 0.8:0.8; 0.7:0.7).

Fig 6-4 CEAC (with variations in the sensitivity and specificity of the HT-test at a fixed price of $4.00) at different WTP values (Replacement role: the HT-test vs. Widal test)



Curves for the other sensitivity and specificity pairs are not shown because at a price of $4, only the (1,1) pair has any chance of being cost-effective, and even then, only above the current cost/QALY threshold of acceptance.

Fig 6-5 CEAC (with variations in the sensitivity and specificity of the HT-test at a fixed price of $1.00) at different WTP values (Replacement role: the HT-test vs. Widal test)



The curve for the (0.7:0.7) sensitivity and specificity pair is not shown because it never goes above zero

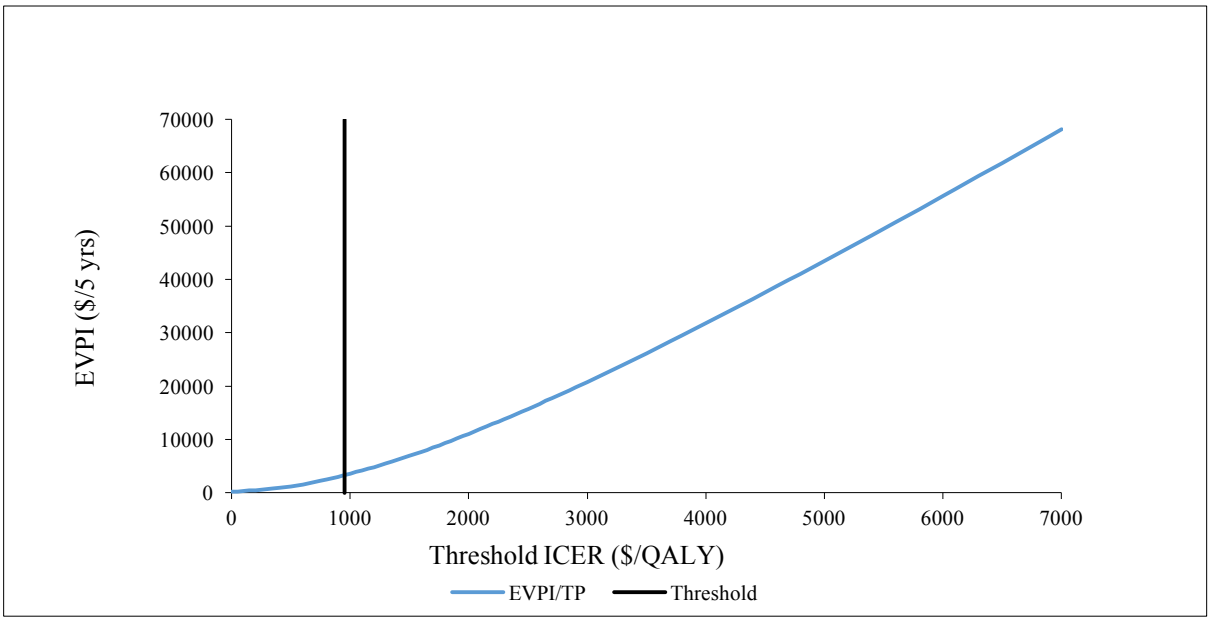
Curves for the (1:1; 0.6:1; 0.9:0.9; 0.7:0.9) sensitivity and specificity pairs are not shown because they never go below one

The results show that specificity is a major driver of the probability of the HT-test being cost-effective at the threshold value of $951 per QALY. For any price above the estimated headroom, the HT-test was always less than 50% likely to be cost-effective at the WTP threshold even for the most optimistic scenario for the HT-test (i.e., both sensitivity and specificity of 100%). For example, the headroom price for the HT-test in the replacement role as shown in Table 6-8 is $3.85. From Fig 6-4, at a price of $4 (which is above the headroom), the HT-test is less than 50% likely to be cost-effective even when both sensitivity and specificity are 100%. The results further show that when the price is low (below the price of the Widal test) (in this case $1), then all the pairs with positive expected incremental effectiveness have a 100% probability of being cost-effective (Figure 6-5).

6.3.1.4 Results for the VOI analysis (the HT-test vs. Widal test)

Using sensitivity and specificity for the HT-test sampled from the same distribution for the Widal tests and a price of $1.00 informed by the CEACs, the results of the VOI analysis for the HT-test vs. Widal test in the replacement role are presented below. The results for the triage and add-on roles are presented in Appendix 9.

Fig 6-6 VOI analysis of the HT-test vs. Widal test in the replacement role at a price of $1.00



As shown in Fig 6-6, at a WTP threshold of $951/QALY, the overall EVPI for the specified decision relevance time horizon of 5 years is $3,287.

In VOI analysis, one may also want information on the parameters causing most of the decision uncertainty and the potential value of reducing the uncertainty by collecting more data. This is known as expected value of partial perfect information (EVPPI). Table 6-9 shows the EVPPI for each of the uncertain model parameters at a WTP threshold of $951/QALY.

Table 6-9 Single parameter EVPPI

Model parameters with uncertainty	Per Person EVPPI ($)	Standard Error	EVPPI for Ghana over 5 years ($)
Mean recovery time for successful typhoid treatment	0	0	0.00
Mean recovery time for successful malaria treatment	0	0	0.00
Prevalence of typhoid in the patient population presenting	0	0	0.00
Probability of first successful typhoid treatment	0	0	0.00
Utility when experiencing typhoid symptoms	0	0	0.00
Specificity of Widal test	0.000222	0	62
Sensitivity of Widal test	0	0	0.00
Specificity of the HT-test	0.001010	0	283
Sensitivity of the HT-test	0	0	0.00

The standard error shows whether there have been enough replications in the model.

As shown in Table 6-9, specificity of the HT-test and Widal test are the two parameters causing most of the decision uncertainty. The standard error values also indicate that in this analysis, there have been enough replications in the model.

Table 6-10 Summary of headroom and percentage likelihood of cost-effectiveness results for the HT-test vs. Widal test

Role	Headroom price ($)	Incremental Effectiveness (QALYs) (positive range)	Price range ($) (across range of positive incremental effectiveness)
Replacement	3.85	0 to 0.0010 (Specificity at least 70%)	1.64-3.85 (Specificity at least 70%)
Triage	1.94	0 to 0.0003 (Specificity at least 90%)	0.34-1.94 (Specificity at least 90%)
Add-on	1.47	0 to 0.0002 (Specificity at least 90%)	0.76-1.47 (Specificity at least 90%)

Role	Price ($)	Percentage likelihood of being cost-effective at threshold for the following accuracy pairs (sensitivity, specificity)							
Replacement		(1,1)	(0.6,1)	(1,0.6)	(0.9,0.9)	(0.7,0.9)	(0.9,0.7)	(0.8,0.8)	(0.70.7)
	4.00	0%							
	3.00	72%	0%	0%	0%	0%	0%	0%	0%
	2.00	100%	63%	0%	89%	7%	0%	0%	0%
	1.00	100%	100%	2%	100%	100%	19%	95%	0%
Triage	2.00	3%							
	1.00	100%	63%	0%	52%	7%	0%	0%	0%
	0.50	100%	100%	1%	99%	91%	7%	33%	0%
Add-on	2	2%							
	1.00	76%	11%	0%	5%	0%	0%	0%	0%
	0.50	100%	91%	0%	34%	10%	0%	1%	0%

Table 6-11 Summary of VOI analysis for the HT-test vs. Widal test

Role	Price ($)	EVPI per person ($)	Overall EVPI ($) for the decision relevant time horizon `(5yrs)
Replacement	2.00	0.00	0.00
	1.00	0.011723	3287
Triage	1.00	0.001029	289
	0.50	0.066283	18585
Add-on	1.00	0.000082	23
	0.50	0.001449	406

Role	Price ($)	EVPPI per person ($) (overall EVPPI for 5 years)($)								
		Mean recovery time for successful typhoid treatment	Mean recovery time for successful malaria treatment	Prevalence of typhoid in the patient population presenting	Probability of first successful typhoid treatment	Utility when experiencing typhoid symptoms	Specificity of Widal test	Sensitivity of Widal test	Specificity of the HT-test	Sensitivity of the HT-test
Replacement	2.00									
	1.00						0.000222 (62)		0.001010 (283)	
Triage	1.00			0.000109 (30.68)						
	0.50			0.05 (14019)		0.000024 (7)	0.000040 (11)	0.000265 (74)	0.010029 (2812)	
Add-on	1.00									
	0.50			0.000307 (86)						

Table 6-10 shows the percentage likelihood of the HT-test being cost-effective compared to the Widal test at a given price and accuracy pair for the HT-test at the WTP threshold value of $951 per QALY. For example, it can be seen from the table that at a price of $2 and test accuracy parameters of sensitivity 60% and specificity 100%, the HT-test is 63% likely to be cost-effective in a replacement role compared to the Widal test. The results also show that, as the price of the HT-test drops, more and more pairs of sensitivity and specificity values become cost-effective.

Table 6-11 shows that for the assumptions made about the HT-test, the EVPI per person is $0.00 when the HT-test is $2.00, thus, the overall EVPI for 5 years is also 0.00. This indicates that research into other parameters to resolve uncertainty is not worth doing based on the assumptions made about the HT-test. The EVPPI results also shows that at this price ($2.00), there is no uncertainty associated with the individual model parameters as an inevitable consequence of the overall EVPI being zero.

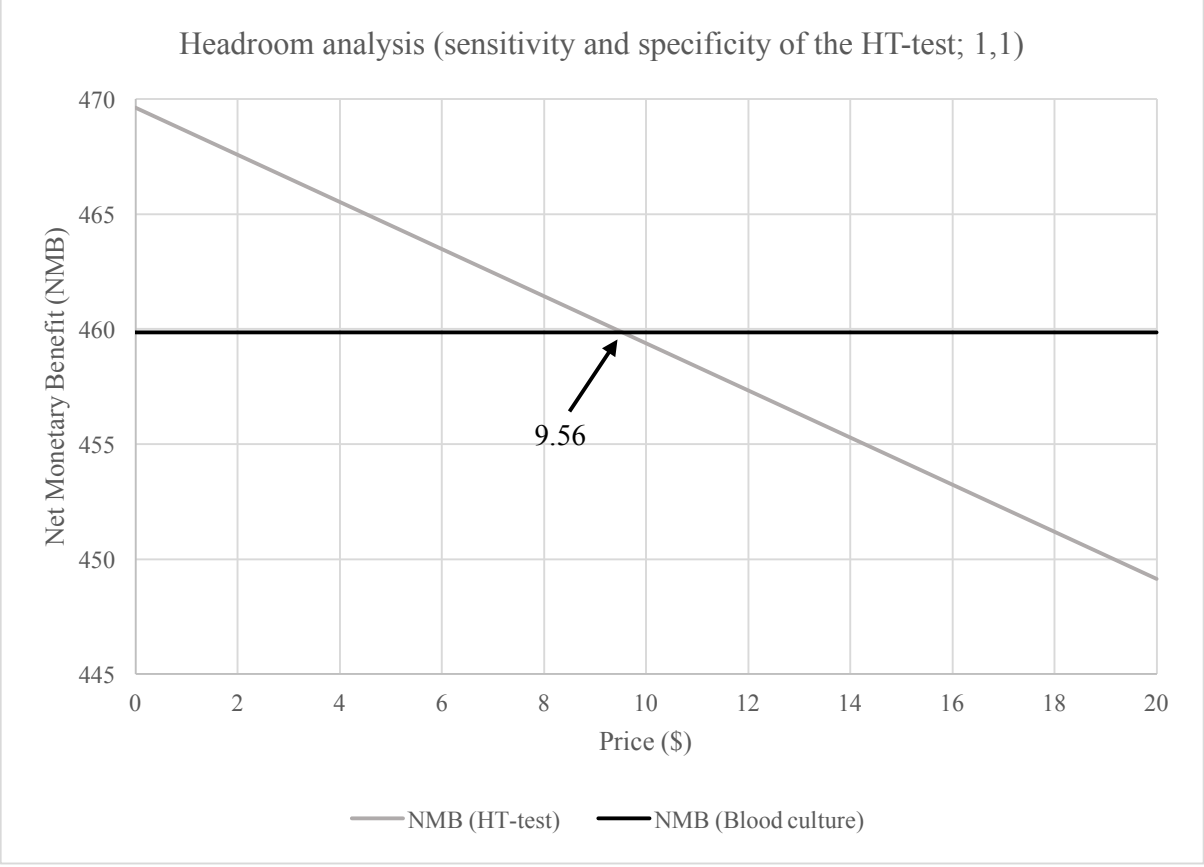
The results further show that, regardless of the scenario, mean recovery time for successful typhoid treatment, mean recovery time for successful malaria treatment, probability of successful first typhoid treatment and the sensitivity of the HT-test do not contribute to the decision uncertainty. Overall, the prevalence of typhoid in the patient population, the specificity of the Widal test and the specificity of the HT-test can be seen to be the parameters contributing to most of the decision uncertainty in the HT-test vs Widal test comparison.

6.3.2 Comparisons of the HT-test vs. blood culture

6.3.2.1 Headroom analysis using NMB (the HT-test vs. blood culture)

Taking all the parameters at baseline and using NMB for each intervention at a WTP for a QALY of $951, the NMB for the HT-test and blood culture with variation in the price for the HT-test in the replacement role is shown in Fig 6-7. As shown in Fig 6-7, at a WTP of $951 for a QALY, the HT-test can be priced up to $9.56 and still be cost-effective vs. blood culture in the replacement role. The maximum prices at which the HT-test can be priced and still be cost-effective vs. blood culture in the triage and add-on roles are $7.14 and $10.08 respectively. The graphs for the triage and add-on roles are presented in Appendix 10.

Fig 6-7 Headroom of the HT-test vs. blood culture at a WTP of $951/QALY (Replacement role)



6.3.2.2 Incremental effectiveness (QALYs) varying sensitivity and specificity pairs for the HT-test vs. blood culture (Replacement role)

The following set of tables show the incremental effectiveness and the maximum price at which the HT-test is still cost-effective (at $951/QALY) at each sensitivity and specificity pair for the HT-test vs. blood culture in the replacement role. The set of tables for the triage and add-on roles are presented in Appendix 11.

The results show a similar observation made in the comparison between the HT-test and Widal test. That is, as the accuracy of the HT-test decreases, the maximum price of the HT-test for it to be cost-effective against blood culture reduces with a decrease in incremental effectiveness in all the roles.

Table 6-12 Incremental effectiveness (QALYs) at each sensitivity and specificity pair for the HT-test vs. blood culture (Replacement role)

Sensitivity \ Specificity	100	90	80	70	60	50	40	30	20	10	0
100	0.0010	0.0007	0.0004	0	-0.0005	-0.0010	-0.0016	-0.0022	-0.0029	-0.0036	
90	0.0010	0.0007	0.0003	-0.0001	-0.0005	-0.0011	-0.0016	-0.0022	-0.0029		
80	0.0009	0.0006	0.0003	-0.0001	-0.0006	-0.0011	-0.0017	-0.0023			
70	0.0009	0.0006	0.0002	-0.0002	-0.0007	-0.0012	-0.0017				
60	0.0008	0.0005	0.0002	-0.0003	-0.0007	-0.0012					
50	0.0007	0.0004	0.0001	-0.0003	-0.0008						
40	0.0007	0.0004	0	-0.0004							
30	0.0006	0.0003	0								
20	0.0006	0.0003									
10	0.0005										
0											

Table 6-13 Maximum ($) price at which the HT-test is still cost-effective at each sensitivity and specificity pair for the HT-test vs. blood culture (Replacement role)

Sensitivity \ Specificity	100	90	80	70	60	50	40	30	20	10	0
100	**9.56**	**8.57**	**7.58**	**6.59**	*5.6*	*4.62*	*3.63*	*2.64*	*1.65*	*0.66*	
90	**9.31**	**8.33**	**7.35**	*6.37*	*5.39*	*4.41*	*3.43*	*2.45*	*1.46*		
80	**9.07**	**8.1**	**7.13**	*6.16*	*5.18*	*4.21*	*3.23*	*2.26*			
70	**8.82**	**7.86**	**6.9**	*5.94*	*4.97*	*4.01*	*3.04*				
60	**8.58**	**7.63**	**6.68**	*5.72*	*4.76*	*3.8*					
50	**8.34**	**7.4**	**6.46**	*5.51*	*4.56*						
40	**8.1**	**7.17**	**6.24**	*5.3*							
30	**7.87**	**6.95**	**6.02**								
20	**7.63**	**6.72**									
10	**7.4**										
0											

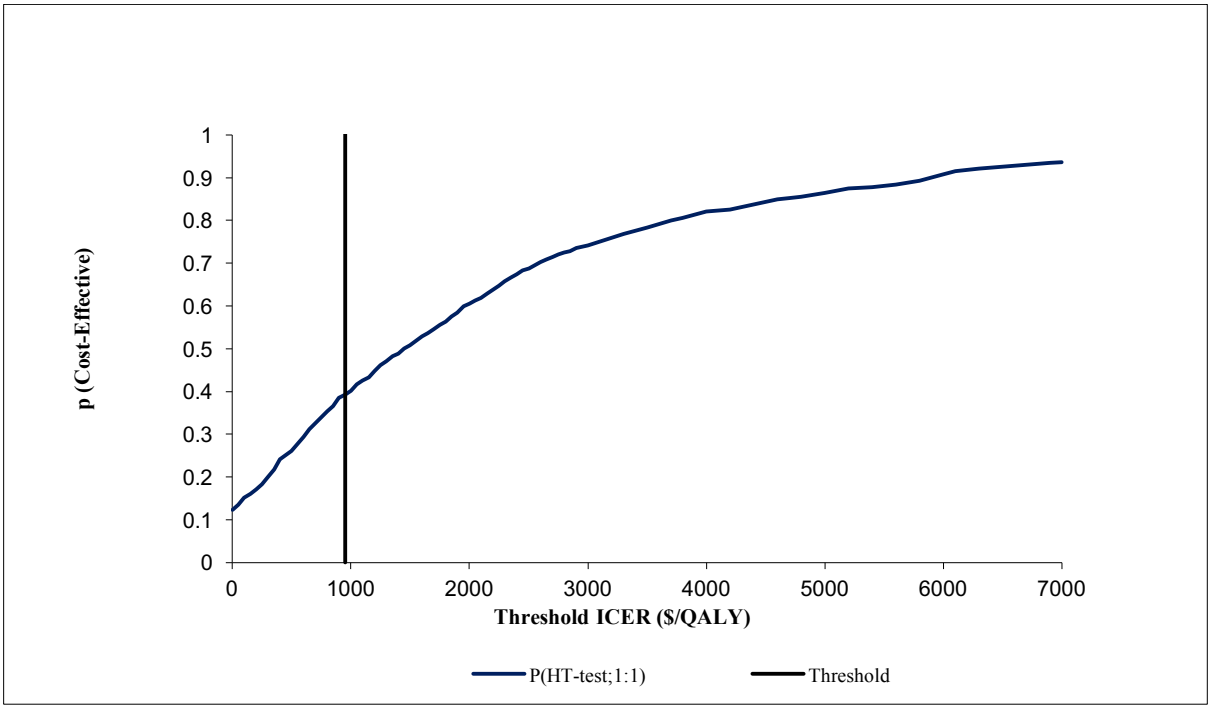
Values in **bold** format represents values of price at sensitivity and specificity pairs at which HT-test is still cost-effective and has a positive incremental effect and would normally be of interest to a decision maker.

Values in *italic* format represent values of price at sensitivity and specificity pairs at which the HT-test is still cost-effective but is slightly less effective than blood culture.

6.3.2.3 Cost-effectiveness acceptability curves for the HT-test vs. blood culture (replacement role)

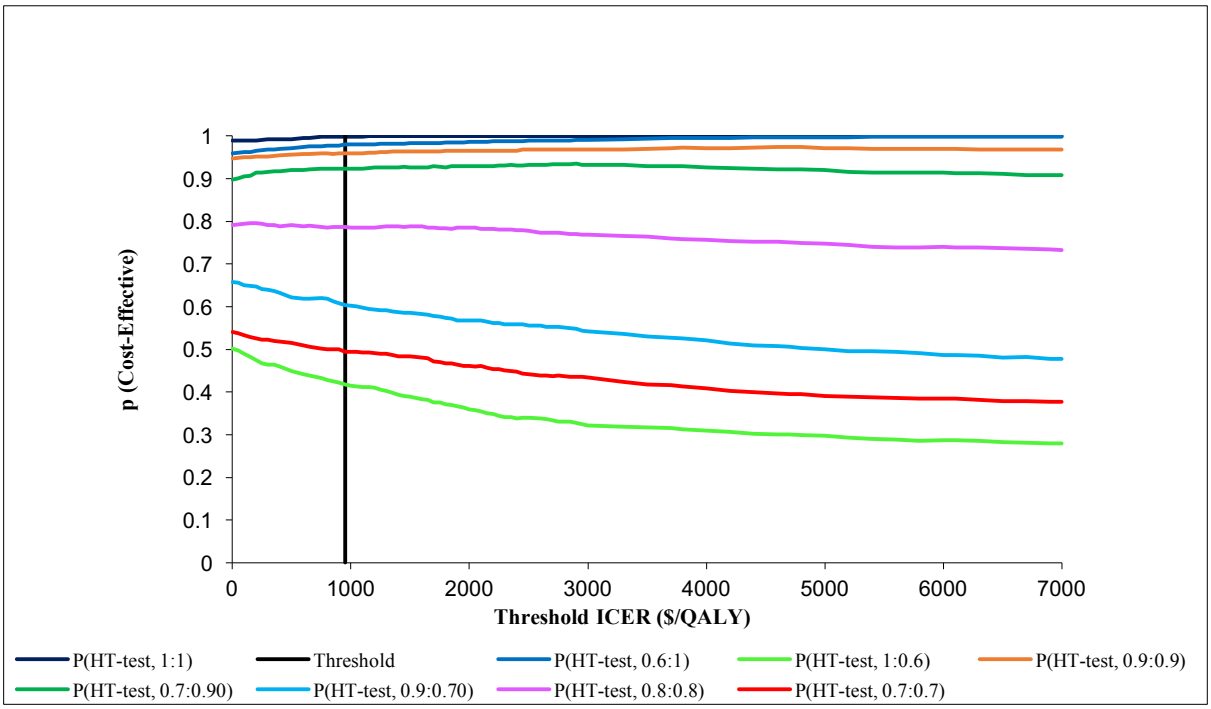
Using a reasonable price range for the HT-test of $6.00 to $10.00, which was informed by findings shown in Table 6-12, and Table 6-13, the CEACs for the PSA of the HT-test vs. blood culture in the replacement role are presented below. The set of graphs for the triage and add-on roles are presented in Appendix 12. In the CEACs, the price is fixed whereas sensitivity and specificity are varied using the following pairs of values (1:1; 0.6:1; 1:0.6; 0.9:0.9; 0.7:0.9; 0.9:0.7; 0.8:0.8; 0.7:0.7).

Fig 6-8 CEAC (with variations in the sensitivity and specificity of the HT-test at a fixed price of $10.00) at different WTP values (Replacement role: HT-test vs. blood culture)



Curves for the other sensitivity and specificity pairs are not shown because at a price of $10.00, only the (1,1) pair has any chance of being cost-effective, and even then, only above the current cost/QALY threshold of acceptance.

Fig 6-9 CEAC (with variations in the sensitivity and specificity of the HT-test at a fixed price of $6.00) at a WTP of $951 (Replacement role: HT-test vs. blood culture)

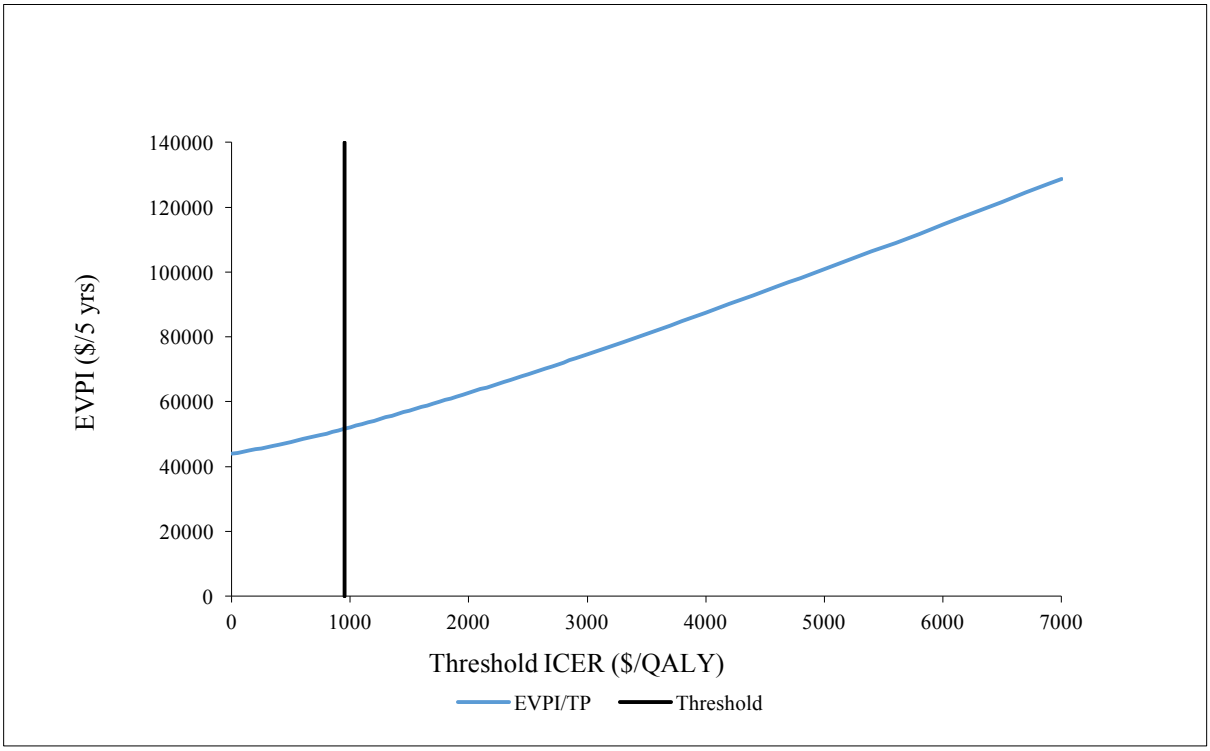


The results showed a similar trend as observed with the HT-test vs. Widal test comparison. That is, specificity is a major driver of the probability of HT-test being cost-effective at the threshold value for cost-effectiveness of $951 per QALY. For any price above the estimated headroom, the HT-test was always less than 50% likely to be cost-effective at the WTP threshold even for the most optimistic scenario for the HT-test (i.e., both sensitivity and specificity of 100%) (Fig 6-8). The results also show that as the price drops, more and more pairs of sensitivity and specificity values become cost-effective (Fig 6-9).

6.3.2.4 Results for the VOI analysis (the HT-test vs. blood culture)

Using sensitivity and specificity for the HT-test sampled from the same distribution for the Widal tests and a price of $6.00 informed by the CEACs, the results of the VOI analysis for the HT-test vs. blood culture in the replacement role are presented below. The results for the triage and add-on roles are presented in Appendix 13.

Fig 6-10 VOI analysis of the HT-test vs. blood culture in the replacement role at a price of $6.00



As shown in Fig 6-10, at a WTP threshold of $951/QALY, the overall EVPI for the specified decision relevance time horizon of 5 years is $51,615.

Table 6-14 shows the EVPPI for each of the uncertain model parameters at a WTP threshold of $951/QALY.

Table 6-14 Single parameter EVPPI

Model parameters with uncertainty	Per Person EVPPI ($)	Standard Error	EVPPI for Ghana over 5 years ($)
Mean recovery time for successful typhoid treatment	0.00	0.00	0
Mean recovery time for successful malaria treatment	0.00	0.00	0
Prevalence of typhoid in the patient population presenting	0.005032	0.00	1411
Probability of first successful typhoid treatment	0.005176	0.00	1451
Utility when experiencing typhoid symptoms	0.00	0.00	0
Specificity of blood culture	0.094824	0.01	26587
Sensitivity of blood culture	0.061716	0.01	17304
Specificity of the HT-test	0.00	0.00	0
Sensitivity of the HT-test	0.00	0.00	0

The standard error shows whether there have been enough replications in the model.

As shown in table 6-14, the parameters causing most of the decision uncertainty are prevalence of typhoid in the patient population, probability of first successful typhoid treatment, sensitivity and specificity of blood culture.

Table 6-15 Summary of headroom and percentage likelihood of cost-effectiveness results for the HT-test vs. blood culture

Role	Headroom price ($)	Incremental Effectiveness (QALYs) (positive range)	Price range ($) (across range of positive incremental effectiveness)
Replacement	9.56	0 to 0.0010 (Specificity at least 70%)	6.59-9.56 (Specificity at least 70%)
Triage	7.14	0 to 0.0003 (Specificity at least 90%)	5.06-7.14 (Specificity at least 90%)
Add-on	10.08	0 to 0.0002 (Specificity at least 90%)	9.37-10.08 (Specificity at least 90%)

Role	Price ($)	Percentage likelihood of being cost-effective at threshold for the following accuracy pairs (sensitivity, specificity)							
Replacement		(1,1)	(0.6,1)	(1,0.6)	(0.9,0.9)	(0.7,0.9)	(0.9,0.7)	(0.8,0.8)	(0.70.7)
	10.00	39%							
	9.00	64%	39%	1%	33%	23%	4%	11%	2%
	8.00	86%	65%	6%	59%	46%	15%	29%	8%
	7.00	97%	88%	20%	83%	74%	35%	54%	25%
	6.00	100%	98%	42%	96%	92%	60%	79%	49%
Triage	8.00	26%							
	7.00	55%	31%	4%	30%	22%	7%	13%	4%
	6.00	83%	64%	14%	61%	49%	24%	35%	18%
	5.00	97%	88%	34%	87%	79%	50%	65%	40%
Add-on	11.00	40%							
	10.00	51%	45%	24%	41%	39%	28%	32%	26%
	9.00	63%	55%	32%	53%	50%	37%	43%	34%

Table 6-16 Summary of VOI analysis for the HT-test vs. blood culture

Role	Price ($)	EVPI per person ($)	Overall EVPI ($) for the decision relevant time horizon `(5yrs)
Replacement	7.00	0.561838	157531
	6.00	0.184085	51615
Triage	6.00	0.434218	121748
	5.00	0.337717	94691
Add-on	10.00	0.546230	153154
	9.00	0.723394	202829

Role	Price ($)	EVPPI per person($) (overall EVPPI for 5 years) ($)								
		Mean recovery time for successful typhoid treatment	Mean recovery time for successful malaria treatment	Prevalence of typhoid in the patient population presenting	Probability of first successful typhoid treatment	Utility when experiencing typhoid symptoms	Specificity of blood culture	Sensitivity of blood culture	Specificity of the HT-test	Sensitivity of the HT-test
Replacement	7.00			0.150315 (42146)	0.072177 (20237)		0.470341 (131876)	0.326644 (91586)	0.003495 (980)	0.000453 (127)
	6.00			0.005032 (1411)	0.005176 (1451)		0.094824 (26587)	0.061716 (17304)		
Triage	6.00			0.236156 (66214)	0.006789 (1903)		0.263581 (73904)	0.186888 (52400)		
	5.00			0.161138 (45180)	0.027984 (7846)		0.142465 (39945)	0.118374 (33190)		
Add-on	10.00			0.234678 (65800)			0.331406 (92921)	0.312028 (87488)		
	9.00			0.412698 (115714)	0.044505 (12479)		0.494482 (138645)	0.476746 (133672)		

Table 6-15 shows the percentage likelihood of the HT-test being cost-effective compared to blood culture at a given price and accuracy pair for the HT-test at the WTP threshold value of $951 per QALY. For example, it can be seen from the table that at a price of $6.00 and test accuracy parameters of sensitivity 90% and specificity 90%, the HT-test is 96% likely to be cost-effective in a replacement role compared to blood culture. The results also show that, as the price of the HT-test drops, more and more pairs of sensitivity and specificity become cost-effective.

Table 6-16 shows that regardless of the scenario, mean recovery time for successful typhoid treatment, mean recovery time for successful malaria treatment and utility when experiencing typhoid symptoms do not contribute to the decision uncertainty. Overall, the prevalence of typhoid in the patient population presenting, probability of first successful typhoid treatment, specificity and sensitivity of blood culture seem to be the parameters contributing to most of the decision uncertainty in the HT-test vs blood culture comparison.

6.4 Discussion

Using a model-based economic evaluation with the outcome measure of the QALY, this chapter examined the potential cost-effectiveness of a hypothetical typhoid test (the HT-test) in Ghana. In this early cost-effectiveness analysis, the HT-test was the index test (the test to be evaluated). The comparator interventions (in this case the reference standards) against which the performance of the HT-test was compared were the culture testing pathway and the Widal testing pathway identified as being the main diagnostic tests used for typhoid fever in Ghana. The evaluation was conducted by varying the position of the HT-test on each of the two testing pathways.

The headroom analysis (where the assumption is made that the HT-test is 100% accurate) shows that the HT-test can be priced up to $3.85 and still be cost-effective against the Widal test in the replacement role. The maximum prices at which the HT-test can be priced and still be cost-effective against the Widal test in the triage and add-on roles are $1.94 and $1.47 respectively. The implication of these results is that depending on the number of tests that are likely to be administered over the lifetime of the HT-test, if it is not possible to develop the HT-test so that it can be sold below these prices and make a profit, then resources should not be committed to its further development. It is worth stating that although a higher price allows for a greater chance of a return on investment, it also increases the possibility of rejection on budget impact grounds.

Examining the headroom room prices for all the different roles together also gives an indication about the likely placement of the HT-test on the current testing pathway. For example, looking at the values $3.85, $1.94 and $1.47, which represents replacement, triage and add-on respectively, it can be concluded that the test has more value in a replacement role rather than

as a triage or add-on to the current testing pathway. That is, if realistically the headroom prices for the triage and add-on roles are too low and it not possible to develop and sell the HT-test at these prices, then this indicates that producing the test to replace the Widal test might be the more viable option. The same principle of reasoning as already described can be applied to the HT-test vs. blood culture comparison. In this comparison, a similar conclusion is made that the HT-test is likely to have more value in a replacement role rather than as a triage or add-on to blood culture.

Acknowledging that the headroom price (when the HT-test is 100% accurate) is probably going to be unrealistic, the focus turns to how accurate the HT-test needs to be to be cost-effective. Additional information on the value of the HT-test compared to each of the comparator tests and in each of the roles was considered based on effectiveness (QALYs) and headroom prices. These results provide provide insights into how good the HT-test needs to be (in terms of accuracy), and at what price, for it to be cost-effective. For example, it is shown that the HT-test can be priced up to $3.85 and still be cost effective against the Widal test in the replacement role when the HT-test has sensitivity and specificity of 100%. In this case, the incremental effectiveness is 0.0010. When the HT-test has a sensitivity and specificity of 70%, the maximum price at which it can still be cost-effective is $0.95 and at an incremental effectiveness of -0.0002. This indicates that as the accuracy of the HT-test decreases, the maximum price at which the HT-test can be priced and still be cost-effective against the Widal test in the replacement role must also decrease. This implies that greater potential returns on investment are possible with a better test. Although care must be taken when recommending a test with characteristics that lead to a negative effectiveness, as further factors beyond cost-effectiveness may need to be considered for the test to be acceptable to decision makers (e.g., small decrease in effectiveness coupled with large cost savings). The maximum price at which

the test can be sold at post development when both sensitivity and specificity are 100% is far greater and more attractive to investors compared to when both parameters are 70% ($3.85 compared to $0.95) and the incremental effectiveness is also more attractive from the patient perspective (i.e., 0.0010 compared to -0.0002). Similar conclusions can be drawn for the other roles of the HT-test vs. Widal test and for the HT-test vs. blood culture.

The PSA revealed the impact of the uncertainty in the parameters on the model results. This shows the probability of the HT-test being cost-effective at different accuracy pairs at a WTP of $951/QALY. The PSA shows that as the price of the HT-test drops, more and more pairs of sensitivity and specificity become cost-effective (i.e., the PSA supports the main findings). The PSA results help to provide a greater understanding of a realistic price and test accuracy parameters for the HT-test to be cost-effective. For example, in the comparison between the HT-test and Widal test (in the replacement role) the headroom price is $3.85, however, at this price the PSA reveals that the HT-test is likely to be cost-effective only when both sensitivity and specificity are 100%. Acknowledging the fact that no test is likely to be 100% accurate, this indicates that it is not realistic for investors to invest in the development of a new test at this price and accuracy pairs.

The PSA results also showed that specificity (at least 70%) is a major driver of the probability of the HT-test being cost-effective at the threshold value for cost-effectiveness of $951 per QALY. This observation is explained by the prevalence of typhoid in the patient population and what happens to FPs and FNs following the test-treat pathway for a suspected typhoid case as shown in the different decision trees. From the models, being FP is worse for patients than being FN because of the extra testing and treatments they undergo before they are eventually cured. Thus, a high specificity is required in order to avoid the unnecessary testing and

treatment associated with these pathways. This observation was common to the analyses for both pathways (i.e., the HT-test vs. Widal test and the HT-test vs. blood culture).

It is shown in the VOI analysis (the HT-test vs. Widal test in the replacement role) that for the specified decision relevance time horizon of 5 years, and for the annual incidence of 60, 000 used in this analysis, the overall expected value of removing decision uncertainty when the HT-test is $1.00 would in total be $3,287 for the assumptions made about the HT-test. This implies that research or data collection exercises costing more than this amount would not be a cost-effective use of resources to resolve this decision uncertainty. This is because the cost of making a wrong decision by decision maker as measured by the cost savings of enabling decision-maker's ability to switch and select other strategies when evidence obtained reduces decision uncertainty, is expected to be no higher than the figure of $3,287. From this value, it can be concluded that there might be no value in resolving this uncertainty as any research studies on further data collection are likely to cost more that this amount.

The VOI analysis further shows that depending on the comparison (i.e., the HT-test vs. Widal test or the HT-test vs. blood culture) different parameters contribute to the decision uncertainty. For example, in the HT-test vs. Widal test comparison, the prevalence of typhoid in the patient population, the specificity of the Widal test and the specificity of the HT-test are the parameters seen to contribute to the decision uncertainty. Whereas in the HT-test vs. blood culture comparison, the prevalence of typhoid in the patient population, probability of first successful typhoid treatment, specificity and sensitivity of blood culture are the parameters seen to contribute to the decision uncertainty. This implies that the focus of any further research to resolve uncertainty in the two comparisons will differ. These results will be of interest to a technology developer as it shows where future research should be focussed.

It is worth mentioning that in this analysis it was assumed that all patients are treated successfully because of the high efficacy of the treatments for typhoid fever and malaria. However, a further consideration in this analysis could have been the inclusion of mortality effects in the model. The implications of including such an effect in the model is that, it may have led to an opposite conclusion as to what was drawn in this study. That is, it may have led to the conclusion that sensitivity is more important than specificity. Furthermore, in this model a value of 1 was assigned to perfect health utility when estimating total QALYs for each branch. Ideally, the baseline health state utility values would be derived from people without specific condition(s) using definitions of health states in the model. However, these data are rarely available (Ara and Brazier, 2011) (as was the case in this analysis), thus, a baseline of full health was assumed. But, because the average person still has other health problems, this assumption (i.e., perfect health utility = 1) may overestimate the benefits of an intervention (in this case, the true value of the HT-test) and thus distort policy decisions (Ara and Brazier, 2011).

An important issue that has been highlighted in this thesis is the inclusion of issues relevant to the patient on the test-treat pathway. Although it is acknowledged that capturing this may have a significant impact on the conclusions drawn from models, in this analysis, the patient's perspective was not captured. This was due to resource constraints (time and money), lack of clear guidance on how this could be done and the methodological complexities associated with capturing these effects (Lee et al., 2010). This highlights the need for concerted efforts to develop methodologies in this field if such effects are to be appropriately measured and valued to be included in economic evaluations. Plausible ways of doing this may include adapting or developing patient preference elicitation methods to collect as much information of users' preference as possible.

6.5 Strengths and limitations of the study

The strengths of this analysis are that extensive sensitivity analysis has been conducted to identify the impact of key parameters on the conclusions drawn from the models. It also highlights the usefulness of early cost-effectiveness modelling embedded within the test development pathway. Furthermore, the models used in this analysis were informed by a formal qualitative approach (different from what normally happens), thus, potentially improving the validity as well as the generalizability and the credibility of the models. The models developed in this study are the first of their type.

A limitation of this analysis is that there was no specific guidance on the best approach to VOI analysis at such an early stage of development to be adopted and how to manage parameters when they are completely unknown and hypothetical. Thus, the approach to VOI analysis adopted in this study may not be the most appropriate. Furthermore, the assumption that patients experience the same utility at the different phases of treatment but return to perfect health once treatment has been completed might be an oversimplification. The assumption of a perfect health utility value of 1 may overestimate the true value of the HT-test and the exclusion of mortality effects from the model may have led to an opposite conclusion being drawn from the analysis. Another limitation of this study is the completely hypothetical nature of the test. Having some form of data on the new test would have been far better than not having any data.

6.6 Conclusion

This evaluation has presented a complete analysis of the potential benefits of the HT-test compared to two alternative tests and in each of three different roles. The analysis shows that

at certain prices and accuracy pairs for the HT-test, the HT-test has potential to be more effective compared to current tests in terms of QALYs gained and cost-effectiveness when introduced into current practice in Ghana. A high specificity (at least 70%) is a key test characteristic requirement for the HT-test to be able to improve current practice. In both pathways, the HT-test is likely to improve current practice when it used to replace the currently used tests.

Chapter 7: Discussion

This chapter discusses the findings of the thesis, its relevance as a valuable source of information to policy makers in Ghana and technology developers, and its wider implications for the early economic evaluation of medical tests.

7.1 Background

The need for the development of a rapid diagnostic test for typhoid fever in Ghana has never been greater because of the serious implications it has both for the health of Ghanaians and on the economy of Ghana. The overall aim of this thesis was to identify the test characteristics of a hypothetical typhoid test such that it would meet cost-effectiveness criteria in the Ghanaian context. To achieve this aim, the thesis addressed four research objectives that are presented in Chapters 3, 4, 5 and 6. The aim of this chapter is to bring together the methodological and empirical work contained in this thesis. The first section provides and overview of the key findings from the study. The next section then considers the strengths and limitations of the thesis. Finally, implications of the findings for policy and future research are presented.

7.2 Issues addressed in this thesis

The ensuing sections provide an overview of the issues addressed in this thesis.

7.2.1 How have economic evaluations conducted at the early phases of test development been done to date?

Whilst several studies have indicated the importance of an iterative use of economic evaluations during the early phases of development of medical devices to support and guide decision-making under conditions of high uncertainty, there is little specific guidance on their

implementation (Buisman et al., 2016). It was found in Chapter 3 that despite this increased interest, there are few published studies that describe the early economic evaluation of medical tests and there is a general lack of clarity on the methods used. Possible explanations for the few studies published in this area include the following: The nature of early economic analysis (i.e., they are primarily used for internal decision-making) implies that this is what makes them less likely to be published. Perhaps because early economic evaluations are not as robust as economic evaluations conducted at the later stages of technology development, that is why they are not considered publishable. Furthermore, test-treat RCTs capturing the downstream health outcomes arising from decisions made based on the test results are not feasible and are rare, thus, the economic evaluation of medical tests involve extensive modelling. This is a difficult thing to do for an established test, but for a new test, it is an even more complex and difficult undertaking as care pathways may not be clearly defined at that early stage of analysis and there is a lack of data.

The key issue emerging from Chapter 3 is that, the tools available to modellers that can demonstrate the value of conducting further research and product development (i.e., VOI analysis, headroom analysis) should be better utilized. The findings of the review conducted here highlights the need for concerted efforts to develop rigorous methodology in this growing field to maximize the value and quality of such analysis. With the limited guidance available on the implementation of early cost-effectiveness analysis, it is expected that the lack of clarity on the appropriate methods to use will continue. However, with the publication of more studies highlighting the need for the development of rigorous methodology in this growing field, there will be an increased awareness on the subject. This is expected to lead to an increase in the effort to develop specific guidance for the proper conduct of such analysis.

7.2.2 Are there any variations in existing test-treat pathways for typhoid fever diagnosis in Ghana?

It was found from Chapter 4 that there is so much variation in practice in terms of typhoid testing and treatment in Ghana. The issue here is that, the quality of care that patients receive is likely to differ depending on the setting in which they find themselves. This situation highlights the need for standardisation of the the test-treat protocol for typhoid fever in Ghana to ensure equitable standards in the quality of care that patients receive regardless of their setting.

Another key issue identified in Chapter 4 was the incorrect use of alternative antimicrobials in addition to the first line treatment of choice for typhoid fever. This situation demonstrates the potential for waste in current practice, highlighting the need for standardisation in, or adherence to, the treatment guidelines for typhoid fever patients. This will promote efficiency in the healthcare delivery system by containing costs associated with the use of additional antimicrobials. Monies saved can be used elsewhere to improve the health of Ghanaians. It will also ensure that patients do not suffer from the side effects associated with unnecessary antimicrobial therapy. Managing these side effects is also associated with a cost, which represents a further waste of limited available resources.

Moreover, there seemed to be a lack of understanding by health professionals (clinicians and medical assistants) in Ghana on test characteristics to make the link between test accuracy and the role of a new test on a care pathway. Clinicians and medical assistants are key stakeholders when delineating test-treat pathways (a key requirement for test evaluation), thus, their lack of understanding of these matters raises issues of concern because this might lead to misinformation on important parameters to capture when modelling test-treat pathways. This

highlights the need for further education on this topic area.

7.2.3 How have economic evaluations focussed on typhoid interventions been modelled?

This issue was addressed in Chapter 5. One of the key findings in this chapter is that there have been relatively few economic evaluations that have focussed on typhoid fever, all of which have focused on the cost-effectiveness of typhoid vaccination.

Another key finding is that transmission dynamic modelling has not yet been incorporated into economic evaluations of typhoid vaccine cost-effectiveness. Thus, there is no economic evaluation of typhoid vaccine cost-effectiveness that will meet the current gold standards for economic evaluation which requires both the direct and indirect benefits of vaccination to be captured when evaluating vaccine cost-effectiveness (Watson and Edmunds, 2015).

7.2.4 What is the potential cost-effectiveness of a hypothetical test for typhoid fever in a Ghanaian context compared to current tests?

The key finding from Chapter 6 is that, the HT-test has potential to perform better than the current available tests in terms of QALYs gained and cost-effectiveness at certain prices and accuracy pairs for the HT-test. However, a high specificity (at least 70%) is required for this to be achieved.

Another key finding as noted from the PSA is that at higher threshold values, the probability of the HT-test being cost-effective increased. However, one of the main issues identified in Chapter 6 is the very low cost-effectiveness threshold used ($951) in this analysis. This greatly

influenced the probability of the HT-test being cost-effective. The implications of this low threshold value to a test developer is that it greatly limits the headroom price for the new technology, thus, there may not be much market potential for new entrants to make profit.

Another key finding from Chapter 6 that confirms the findings from an earlier chapter is that, the techniques for early cost-effectiveness analysis are not well developed. Particularly in this chapter, there was the issue with the most appropriate approach to adopt in VOI analysis. In the traditional VOI analysis, there is always some form of data available to inform estimates of the new technology to enable the assessment of any value in further research. The difficulty in this analysis was that there was no information on the hypothetical test and in the absence of any specific guidance, there was no clear direction as to what to do. It is acknowledged that although the approach adopted in this study (the assumption was made that the HT-test has the same accuracy characteristics as the Widal test) might not be the most appropriate, it certainly highlights an area requiring further research. The difficulty in conducting VOI analysis at the early stages of development could be a possible explanation of their omission from studies as noted in Chapter 3.

7.3 Strengths of the research

One of the major strengths of this thesis is that it systematically reviewed previous early modelling studies and typhoid studies to enable the most robust methodologies to the early modelling approach to be adopted in this thesis.

A formal qualitative methodology was used to capture and analyse the perspectives of 51 participants working in different settings (rather than a few informants) to delineate the existing test-treat pathway for typhoid fever which then assisted in model structure conceptualisation

(i.e., primary data was used to inform the model structures used in the analysis). Capturing variability could potentially assist in improving the validity as well as the generalizability and credibility of the models used in this thesis.

Extensive sensitivity analysis was conducted to identify the the impact of key parameters on the conclusions drawn from the models.

Finally, the thesis also has the strength of highlighting the usefulness of early cost-effectiveness modelling embedded within the test development pathway.

7.4 limitations of the research

Although this thesis has addressed some important issues, like any other study, it has its own limitations.

The lack of a transmission model in this analysis may mean that the impact of a better test may be underestimated. Furthermore, the lack of specific guidance on the best approach to VOI analysis at such an early stage to be adopted for the analysis implies that the approach utilized may not be the most appropriate.

Another limitation is that only the perspectives of clinicians and medical assistants were captured and in only one region when delineating the existing care pathways for typhoid fever in Ghana. If other relevant stakeholders such as patients (van Voorn et al., 2016) and policy makers had been sampled, they may have held different views or added further considerations. Finally, it was not possible to probe the responses of participants (clinicians and medical assistants) because the interviews were not personally conducted. For example, in this thesis,

this was a deficiency when capturing valid participants' views about the anticipated role of a new test.

7.5 Implications for future research

This thesis has shown that the techniques for early economic evaluation are not well developed; particularly, there is no specific guidance on how to conduct VOI analysis at the early stages of test development. This highlights the need for future research in this field to develop robust methodologies if the value of such analysis is to be maximized.

7.6 Relevance of study findings to policy makers in Ghana

The findings from this thesis provide valuable information to policy makers in Ghana to make policy changes that will greatly improve the delivery of care for typhoid fever patients in Ghana.

The study findings showed that some clinicians and medical assistants in Ghana presumptively treat typhoid fever (although in the minority) for valid reasons such as delays in diagnosis or concerns about accuracy of test or lack of availability. These circumstances reflect the characteristics of current practice in Ghana and the implication to policy makers is that, when making the decision to introduce a new test into current practice, the test would need to be "better" in these respects for healthcare professionals in Ghana to use. The new test should be accessible with a quick turnaround time and be at least as accurate as both tests.

Furthermore, the use of alternative antibiotics in addition to ciprofloxacin demonstrates the potential for waste in current practice. The implication to policy makers is that, more education is needed to ensure the standardisation in, or adherence to, the treatment regimen for typhoid fever patients. This will not only promote efficiency in the healthcare delivery system but will

also ensure that, patients do not suffer from the side effects associated with unnecessary antimicrobial therapy (the management of which is also associated with a cost to the healthcare system). Monies saved could be used elsewhere to improve the health of Ghanaians.

The preliminary results of the early cost-effectiveness analysis showed the potential for the HT-test to be more effective compared to current tests in terms of QALYs gained and cost-effectiveness when introduced into current practice in Ghana. The implication to policy makers is that, it provides them with valuable information that will guide their decision to fund the further development of the HT-test. This could help allocate limited budgets more efficiently.

7.7 Relevance of study findings to technology developers

The findings of this thesis provide technology developers with information about the market potential of developing the HT-test. It provides information to the technology developer on the headroom prices and test characteristics at which the HT-test can be produced for it to be cost-effective and be able to make profit on any investments that would be made in the further development of the test. It also provides information to the test developer on overall EVPI and EVPPI to support decision-making on resolving uncertainty and the parameters to focus on.

7.8 Conclusion

The aim of the research undertaken within this thesis was to examine the methodologies and tools that have been employed in the early economic evaluation of medical tests, to explore the variations in existing test-treat strategies for typhoid fever in Ghana to delineate the care pathway for typhoid fever, to explore the types of models that have been adopted for test-treat evaluations of typhoid fever and to design an economic model to estimate the maximum price

and the minimum test performance required for the HT-test to be cost-effective in the Ghanaian setting.

The findings of this research have wider implications for the early economic evaluation of tests. This research has made a number of important contributions to the early economic evaluation literature, both methodologically, and in terms of its recommendations for further research. This thesis therefore provides valuable information relevant to developing guidance in the implementation of early cost-effectiveness analysis. It has been noted that there is a general lack of clarity on the methods used in early economic evaluations, hence, it has been recommended that there is the need to develop robust methodologies in this field if the value of such analysis is to be maximized. This research further recommends that to maximize the value of early economic evaluations, the methods available to modellers that can demonstrate the value of conducting further research and product development (i.e., value of information (VOI) analysis, headroom analysis) should be better utilized. Furthermore, it has been recommended that extensive sensitivity analysis should be incorporated into such analysis because of the nature of early economic evaluations (i.e., associated with limited data and high uncertainty), to make them useful and relevant to policy making. There is no specific guidance on VOI analysis at this early stage of analysis and it is recommended here that, further research is needed in this field.

To the author's knowledge, the models developed in this study are the first of their type. The research has used a robust qualitative approach and demonstrated the benefits associated with integrating formal qualitative methods in the model development process. This thesis therefore provides a rigorous methodology for those wanting to undertake further research to improve

the methods used in modelling processes and/or to develop modelling guidance. It also provides a simple but rigorous methodology that appeals to quantitative researchers who would want to employ formal qualitative approaches in their research.

The contribution of this research is also important in the context of improving typhoid testing and treatment not only in Ghana, but also in other resource-limited settings where this is a problem. It has been demonstrated here that, for a hypothetical test to perform better than the current available tests in Ghana in terms of QALYs gained and cost-effectiveness, it is necessary that it has a high specificity (at least 70%) and should not be priced more than $10. The findings of this thesis also provide valuable information to test developers about the market potential of developing a new test for typhoid fever that can be applied in the Ghanaian setting and the focus of any further research to resolve uncertainty.

APPENDICES

Appendix 1: Search strategy in Medline

Table A1-1 Search strategy in Medline

#	**Searches**	**Results**
1	((new or novel or emerging or innovati* or early) adj4 test*).tw.	43724
2	((earl* or mid* or develop* or formative* or determinative* or design* or concept* or investigation*) adj2 (phase* or stage* or process* or "life cycle" or lifecycle) adj4 test*).tw.	1971
3	((nascent or original or "ground breaking" or groundbreaking or promising or "cutting edge" or "cutting-edge" or seminal or "recently introduced") adj3 test*).tw.	3534
4	1 or 2 or 3	48091
5	(new or innovat* or emerging* or novel* or nascent or original or "ground breaking" or groundbreaking or promising or "cutting edge" or "cutting-edge" or seminal or "recently introduced").tw.	2670140
6	exp "diagnostic test"/	8662
7	exp screening/	109672
8	exp prognosis/	1267953
9	exp "predisposition test"/	30098
10	6 or 7 or 8 or 9	1405426
11	5 and 10	180738
12	(early adj3 (valuation* or evaluation* or assessment* or model*)).tw.	13183
13	(early-stage adj3 (valuation* or evaluation* or assessment* or model*)).tw.	417
14	(early adj2 (economic or estimate*)).tw.	474
15	12 or 13 or 14	13702
16	15 and (4 or 11)	733
17	(headroom or "head room").tw.	34
18	models, economic/ or models, econometric/	11782
19	markov chain/	11322
20	Decision Trees/	9553
21	decision support techniques/	15041
22	(discrete event* adj8 model*).tw.	361
23	(decision* adj5 model*).tw.	9742
24	(model* adj5 markov*).tw.	8787
25	((econom* or cost or costs) adj model*).tw.	2837
26	decision making/	76162
27	cost benefit analysis/	66482
28	technology assessment, biomedical/	8612
29	18 or 19 or 20 or 21 or 22 or 23 or 24 or 25 or 26 or 27 or 28	188046
30	17 and 29	8

31	16 and 29	27
32	11 and 29	4121
33	method*.tw.	4035419
34	32 and 33	2131
35	16 or 30 or 31 or 34	2866

Appendix 2: Typhoid fever diagnosis and treatment questionnaire

The following is the exact text of the questionnaire used

Please take a few minutes of your time to read and fill out this questionnaire to the best of your ability. Your responses will be treated as highly confidential.

1. What are the signs and symptoms typhoid fever patients present with?

2. What are your views on applying presumptive treatment to typhoid fever?

3. What are the test(s) available to you if you wanted to test for typhoid fever before treatment?

4. Which test(s) for patients presenting with symptoms of typhoid fever do you normally use?

5. Why do you use the test(s) you mentioned in 4 above?

6. In the absence of a perfectly accurate test, there always exist a trade-off between the sensitivity and specificity of a test. Which is more important to you: correctly identifying patients with typhoid fever (sensitivity) or avoiding unnecessary typhoid treatment for patients that don't have the disease (specificity) and why?

7. What other properties do you consider important components of a test for the diagnosis of typhoid fever?

8. If there was a new test with the properties you have stated in questions 6 & 7 above, would you still use the test(s) you use now (as stated in question 4) for making clinical management decisions?

9. How do you see yourself using this new test if it was available to you? (please tick)

☐	To test patients first to decide on who should receive further testing with the test you use
☐	As the main investigative tool without any further testing with the test you use
☐	To provide additional diagnostic information after testing with the test you use

10. How do you normally manage typhoid test positive and typhoid test negative OPD patients?

11. What are the options of antimicrobials available to you to decide a treatment plan for typhoid test positive OPD patients?

12. What is the prescription of the antimicrobial you normally give to treat typhoid test positive OPD patients?

13. Do you prescribe the same antimicrobial for every typhoid test positive OPD patient (in terms of brand and class)?

14. What is/are the reason(s) for your answer in question 13 above?

15. How do you assess whether your test-treat strategy has been effective in improving patient health?

Appendix 3: The "analytical framework matrix"

Table A3-1 The "analytical framework matrix"

Participant	Type of facility	Presumptive treatment views	Diagnostic test(s) used	Test (-ve) management	Test (+ve) management	Prescribed antimicrobial for (+ves)	Dosage of antimicrobial	Assessment of test-treat outcome	Intende-d Role of new test	Test property considered important (sensitivity or specificity)
Participant 1	Polyclinic	May lead to treatment failure	Widal test	Health education	Antimicrobial therapy	Oral ciprofloxacin Oral Metronidazole	Tab Ciprofloxscin 500mg bd*7 days Tab Metronidazole 200-400mg tds * 7 days	Review and reassess	Add-on	sensitivity
Participant 2	Primary hospital	very subjective and not evidence based	Widal test	Not answered	Antimicrobial therapy	Oral ciprofloxacin	Not answered	Repeat test at least 10 days after treatment	Add-on	sensitivity
Participant 3	Primary hospital	Useful when diagnosis yet to made	Widal test Blood culture	Further testing for other common ailments	Antimicrobial therapy	Oral Ciprofloxacin	Tab Ciprofloxscin 500mg bd* 7-14 days	Repeat diagnostic test	Add-on	specificity
Participant 4	Primary hospital	Should be used ONLY when test not available	Widal test FBC	Prophylaxis	Antimicrobial therapy	Oral ciprofloxacin Oral Cefixime Inj Ceftriaxone	Tb Ciprofloxacin 500mg bd* 7 days Tb Cefixime 200mg bd* 7 days Inj Ceftriaxone 1.5g stat then 750mg tds*24 hrs	Review and retest	Add-on	specificity
Participant 5	Primary hospital	May be reliable since available Widal test is unreliable	Widal Test	Health education on hygienic practices	Antimicrobial therapy	Oral Ciprofloxacin Oral Azithromycin	Tb Ciprofloxacin 500mg bd* 14 days. Tb Azithromycin 1g daily*7	Repeat test on review	Add-on	sensitivity

Participant 6	Primary hospital	When logistics for testing not available	Widal Test FBC Stool R/E	Prophylaxis	Antimicrobial therapy	Oral Ciprofloxacin Oral Metronidazole Inj Ceftriaxone	Tb Ciprofloxacin 500mg bd* 14 days. Tb Metronidazole 400mg tds*14 days In Ceftriaxone 2g stat	Repeat test on review	Triage	sensitivity
Participant 7	Primary hospital	Used in the absence of test for typhoid	Widal Test	Prophylaxis	Antimicrobial therapy	Oral Ciprofloxacin Inj Ceftriaxone	Tb Ciprofloxacin 500mg bd* 5 days Inj Ceftriaxone 2g Stat, then 1g daily* 48 hrs	Repeat test on review	Add-on	sensitivity
Participant 8	Primary hospital	Not a good practise	Widal Test Stool C/S	Symptomatic treatment	Antimicrobial therapy	Oral Ciprofloxacin Oral Azithromycin	Tab Ciprofloxacin 500mg bd* 14 days Tb Azithromycin 1g daily* 7days	Repeat test on review	Add-on	sensitivity & specificity
Participant 9	Primary hospital	Used in the absence of test for typhoid	Widal Test	Prophylaxis	Antimicrobial therapy	Oral Ciprofloxacin Inj Ceftriaxone	Tb Ciprofloxacin 500mg bd* 7 days Inj Ceftriaxone 2g stat, then 1g daily *48 hrs	Repeat test on review	Triage	sensitivity
Participant 10	Primary hospital	Sometimes fails	Widal Test FBC Stool RE	Prophylaxis	Antimicrobial therapy	Oral Ciprofloxacin Oral Ofloxacin Inj Ceftriaxone/Cef uroxime	Not answered	Repeat test on review	Triage	specificity
Participant 11	Primary hospital	Can cause antimicrobial resistance	Widal Test FBC	Prophylaxis	Antimicrobial therapy	Oral Ciprofloxacin Oral Metronidazole Inj Ceftriaxone	Not answered	Reviewing the patient	Triage	specificity
Participant 12	Primary hospital	Not too presumptive. Testing is encouraged	Widal Test	Test and treat other similar conditions	Antimicrobial therapy	Inj Ceftriaxone Oral Ampicillin Oral Trimethoprim Sulfamethoxazol e	Not answered	Repeat test on review	Triage	sensitivity
Participant 13	Primary hospital	Should be encouraged	Widal Test	Test and treat other conditions	Antimicrobial therapy	Oral Ciprofloxacin	Tb Ciprofloxacin 500mg bd* 14 days	Repeat test on review Sometimes patients verbally say they are better	Add-on	Not answered

Participant 14	Primary hospital	Not professional Can lead to bacteraemia, drug resistance, complications	Widal Test	Not answered	Antimicrobial therapy	Oral Ciprofloxacin Oral Cefuroxime	Tb Ciprofloxacin 500mg bd* 14 days Tb Cefuroxime 250mg bd*7 days	Repeat test on review. Reassess patient's condition	Add-on	sensitivity
Participant 15	Primary hospital	Should be improved	Widal Test	Test and treat other conditions. E.g., malaria, enteritis	Antimicrobial therapy	Oral Ciprofloxacin Oral Azithromycin	Tb Ciprofloxacin 500mg bd* 14 days	Repeat test on review	Add-on	specificity
Participant 16	Primary hospital	Shouldn't be encouraged	Widal Test	Test and treat for other conditions like Malaria	Antimicrobial therapy	Oral Ciprofloxacin Oral Azithromycin	Tb Ciprofloxacin 12hly* 14 days Tb Azithromycin 500mg 12hly* 14 days	Review and reassess with test	Replace-ment	specificity
Participant 17	Primary hospital	Usually essential	Widal Test Stool R/E Blood film for malaria parasites	Further testing and treated for other conditions	Antimicrobial therapy	Oral Ciprofloxacin	Tb Ciprofloxacin 500mg bd*7 days	Repeat test on review	Triage	specificity
Participant 18	Primary hospital	Treatment failure and antimicrobial resistance	Widal Test	Investigate other causes of ailment	Antimicrobial therapy	Oral Ciprofloxacin	Tb Ciprofloxacin 500mg bd* 7 days	Review with assessment	Triage	sensitivity
Participant 19	Primary hospital	Antimicrobial resistance Treatment failure	Stool Culture and sensitivity	Test and treat other similar presenting conditions	Antimicrobial therapy	Oral Ciprofloxacin Oral Cefixime	Tb Ciprofloxacin 500mg bd* 7 days Tb Cefixime 400mg bd* 7 days	Repeat test on review ("patients no longer complain of signs and symptoms")	Add-on	sensitivity
Participant 20	Primary hospital	Hampers thorough treatment	Widal Test	Investigate other causes of ailment and treat appropriately	Antimicrobial therapy	Oral Ciprofloxacin	Tb Ciprofloxacin bd * 7-14 days	Review patients after initial treatment	Triage	sensitivity

Participant 21	Primary hospital	Treatment must be based on clinical predictors	FBC	Not answered	Antimicrobial therapy	Oral Ciprofloxacin	Tb Ciprofloxacin bd * 7 days	Follow-ups	Replace-ment	sensitivity
Participant 22	Primary hospital	Leads to antimicrobial resistance	Widal Test	Investigate other conditions and treat accordingly	Antimicrobial therapy	Oral Ciprofloxacin	Tb Ciprofloxacin bd * 7-14 days	Repeat test on review	Triage	sensitivity
Participant 23	Primary hospital	Treatment can be started if clinical features present. Test available not solely reliable	Widal Test	Test and treat other conditions	Antimicrobial therapy	Oral Ciprofloxacin Oral cefuroxime	Not answered	Follow-ups and re-testing after treatment	Add-on	specificity
Participant 24	Health centre	Antimicrobial resistance	Not answere-d	Not answered	Not answered	Not answered	Not answered	Not answered	Not answere-d	Not answered
Participant 25	Primary hospital	Treat presumptively in severe case if suspected	Blood for C/S	Not answered	Antimicrobial therapy	Oral Ciprofloxacin	Tb Ciprofloxacin 500mg bd* 14 days	Review	Triage	sensitivity
Participant 26	Secondary hospital	Should be avoided; presents similar to several diseases	Blood culture	Not answered	Antimicrobial therapy	Oral Amoxicillin	Not answered	Repeat test on review	Add-on	sensitivity
Participant 27	Secondary hospital	Confirmation before treatment	Blood culture	Investigate for other diseases and treat	Antimicrobial therapy + Treatment for associated signs (e.g., Antipyretic for fever; antimotility for diarrheoa)	Oral Amoxicillin	Not answered	Repeat test on review	Add-on	sensitivity
Participant 28	Secondary hospital	Antimicrobial resistance	Blood culture Stool culture	Not answered	Antimicrobial therapy	Oral Ciprofloxacin Oral Azithromycin	Tb Ciprofloxacin 500mg bd* 14 days	Review with reassessment	Triage	sensitivity

Participant 29	Secondary hospital	Effective; saves time and lives	Stool culture Blood culture	Not answered	Antimicrobial therapy	Oral Ciprofloxacin Oral Levofloxacin	Tb Ciprofloxacin bd * 7-14 days	Repeat test on review	Triage	sensitivity
Participant 30	Secondary hospital	Good to prevent complications ("Since the most precise blood/stool cultures are not available in most districts and health centres")	Urine R/E Blood film for malaria FBC	Further investigation	Antimicrobial therapy	Oral Cefixime	Tb Cefixime 400mg bd* 10-14 days Tb Ciprofloxacin bd * 7-14 days	Full recover on review	Triage	sensitivity & specificity
Participant 31	Secondary hospital	Shouldn't be done	Widal Test	Further investigation and treat	Antimicrobial therapy	Oral Ciprofloxacin	Tb Ciprofloxacin 500mg bd* 7-14 days	Repeat test after completion of treatment	Add-on	sensitivity
Participant 32	Secondary hospital	Not answered	Widal Test Stool culture	Further test and treat appropriately	Antimicrobial therapy	Oral Ciprofloxacin Oral Cefuroxime	Tb Ciprofloxacin 500mg bd* 10 days Tab Cufuroxime 500mg bd* 7 days	Review and reassess	Add-on	sensitivity
Participant 33	Secondary hospital	Should be discouraged; could lead to antimicrobial resistance	Stool culture	Non pharmacological (hygiene advice/counselling)	Antimicrobial therapy + Non pharmacological	Oral Ciprofloxacin	Tb Ciprofloxacin 500mg bd* 14 days	Calling and making enquiries from patient	Replace-ment	sensitivity
Participant 34	Secondary hospital	Don't like the idea. Test before treat	Stool culture Blood culture	I do not treat	Antimicrobial therapy	Oral Ciprofloxacin	Tb Ciprofloxacin 500mg bd* 14 days	Review patient	Add-on	sensitivity
Participant 35	Secondary hospital	Better to test before treatment	Cultures (urine, blood, stool) FBC	Investigate underlying cause of symptoms and treat	Antimicrobial therapy	Oral Ciprofloxacin	Not answered	Repeat test on review	Add-on	sensitivity
Participant 36	Secondary hospital	Misdiagnosis	Widal test Culture (stool,	No treatment	Antimicrobial therapy	Oral Ciprofloxacin	Tb Ciprofloxacin 500mg bd* 7-14 days	Not answered	Replace-ment	specificity

			urine blood)							
Participant 37	Secondary hospital	Diagnosis should always be confirmed before treatment started	Stool/Urine culture FBC	Not answered	Antimicrobial therapy	Oral Ciprofloxacin Oral Azithromycin	Tb Ciprofloxacin 500mg bd* 7-14 days	Retest after treatment	Add-on	specificity
Participant 38	Secondary hospital	When there is high index of suspicion "e.g., someone who often buys food from unhygienic sources"	Stool culture	Investigate further for specific condition and treat	Antimicrobial therapy	Oral Ciprofloxacin	Not answered	Not answered	Replacement	specificity
Participant 39	Secondary hospital	Antimicrobial resistance	Blood culture Stool culture	No need to manage	Antimicrobial therapy	Oral Ciprofloxacin	Tb Ciprofloxacin 500mg bd* 7-14 days	Check if signs and symptoms have been resolved	Add-on	sensitivity
Participant 40	Secondary hospital	Antimicrobial resistance	Blood culture Stool culture	Not answered	Antimicrobial therapy	Oral Ciprofloxacin	Tb Ciprofloxacin 500mg bd* 7-14 days	Review after treatment	Triage	sensitivity
Participant 41	Secondary hospital	Antimicrobial resistance	Blood culture Stool culture	Not answered	Antimicrobial therapy	Oral Ciprofloxacin Oral Azithromycin	Tb Ciprofloxacin 500mg bd* 7-14 days	Review	Triage	sensitivity
Participant 42	Secondary hospital	Antimicrobial resistance	Blood culture Stool culture	Investigate further and treat	Antimicrobial therapy	Oral Ciprofloxacin Oral Cefotaxime	Tb Ciprofloxacin 500mg bd* 7-14 days	Review	Triage	sensitivity
Participant 43	Secondary hospital	Misdiagnosis	Blood culture	Investigate and treat accordingly	Antimicrobial therapy	Oral Ciprofloxacin Oral Metronidazole	Tb Ciprofloxacin 500mg bd* 7days Tb Metronidazole 400mg tds* 7 days	Review with test	Add-on	sensitivity
Participant 44	Secondary hospital	Antibiotic resistance	Cultures (urine, blood, stool)	Not answered	Antimicrobial therapy	Oral Ciprofloxacin	Tb Ciprofloxacin 500mg bd* 7-14 days	Retest after treatment	Replace-ment	sensitivity

Participant 45	Secondary hospital	Misdiagnosis	Blood culture FBC	Test for other conditions and treat accordingly	Antimicrobial therapy	Oral Ciprofloxacin	Tb Ciprofloxacin 500mg bd* 7-14 days	Follow-ups	Add-on	specificity
Participant 46	Primary hospital	Misdiagnosis Treatment failure	Widal test	Test for other conditions and treat	Antimicrobial therapy	Oral Ciprofloxacin	Tb Ciprofloxacin 500mg bd* 7 days	Retest on review	Add-on	sensitivity
participant 47	Primary hospital	Treatment is more of blind treatment	Widal test	Not answered	Antimicrobial therapy	Oral Ciprofloxacin	Tb Ciprofloxacin 500mg bd* 14 days	Retest after treatment	Add-on	specificity
Participant 48	Primary hospital	Misdiagnosis Treatment failure	Widal test	Test for other conditions and treat	Antimicrobial therapy	Oral Ciprofloxacin	Tb Ciprofloxacin 500mg bd* 7-14 days	Retest after treatment	Add-on	sensitivity
Participant 49	Primary hospital	Treatment failure Antimicrobial resistance	Widal test	Education on hygiene	Antimicrobial therapy	Oral Ciprofloxacin IV Ceftriaxone	Tb Ciprofloxacin 500mg bd* 7 days	Repeat test on review	Add-on	specificity
Participant 50	Primary hospital	Can be started since cultures takes 72 hrs	Blood culture Stool culture	Further investigation	Antimicrobial therapy	Oral Ciprofloxacin	Tb Ciprofloxacin 500mg bd* 7-10 days	Repeat test after treatment	Add-on	specificity
Participant 51	Primary hospital	antimicrobial resistance	Widal test	Further investigation	Antimicrobial therapy	Oral ciprofloxacin	Tb Ciprofloxacin 500mg bd* 10 days	Review and repeat testing	Add-on	sensitivity

Appendix 4: Search strategy for each database used in Chapter 5

Database(s): **Ovid MEDLINE(R)**
Search Strategy:

#	Searches	Results
1	Cost-Benefit Analysis/ or economic evaluation*.mp.	74470
2	typhoid fever.mp. or Typhoid Fever/	11571
3	enteric fever.mp. or Typhoid Fever/	10913
4	2 or 3	11819
5	1 and 4	30

Database(s): **Embase**
Search Strategy:

#	Searches	Results
1	"cost effectiveness analysis"/ or economic evaluation/ or economic evaluation*.mp.	141701
2	typhoid fever.mp. or typhoid fever/	12727
3	enteric fever.mp. or typhoid fever/	12126
4	2 or 3	13014
5	1 and 4	95
6	limit 5 to (exclude medline journals and yr="1948 - 2017")	9

Database: **DARE, HTA, NHS EED**

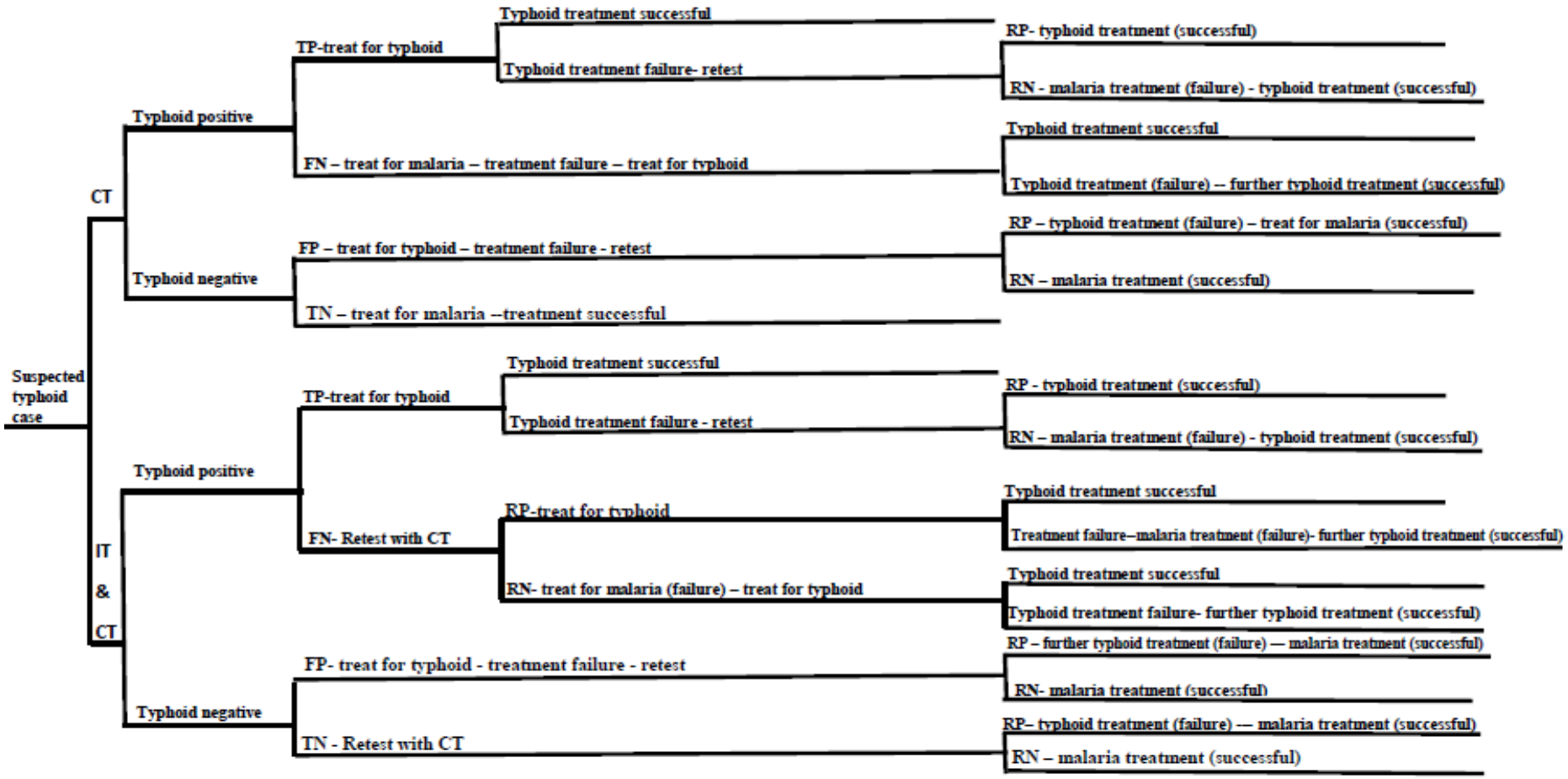
1	(economic evaluation) IN DARE, NHSEED, HTA	18521
2	(typhoid fever) IN DARE, NHSEED, HTA	15
3	#1 AND #2	7

Database: **PubMed**

("typhoid fever"[MeSH Terms] OR ("typhoid"[All Fields] AND "fever"[All Fields]) OR "typhoid fever"[All Fields]) AND ("cost-benefit analysis"[MeSH Terms] OR ("cost-benefit"[All Fields] AND "analysis"[All Fields]) OR "cost-benefit analysis"[All Fields] OR ("economic"[All Fields] AND "evaluations"[All Fields]) OR "economic evaluations"[All Fields])	30

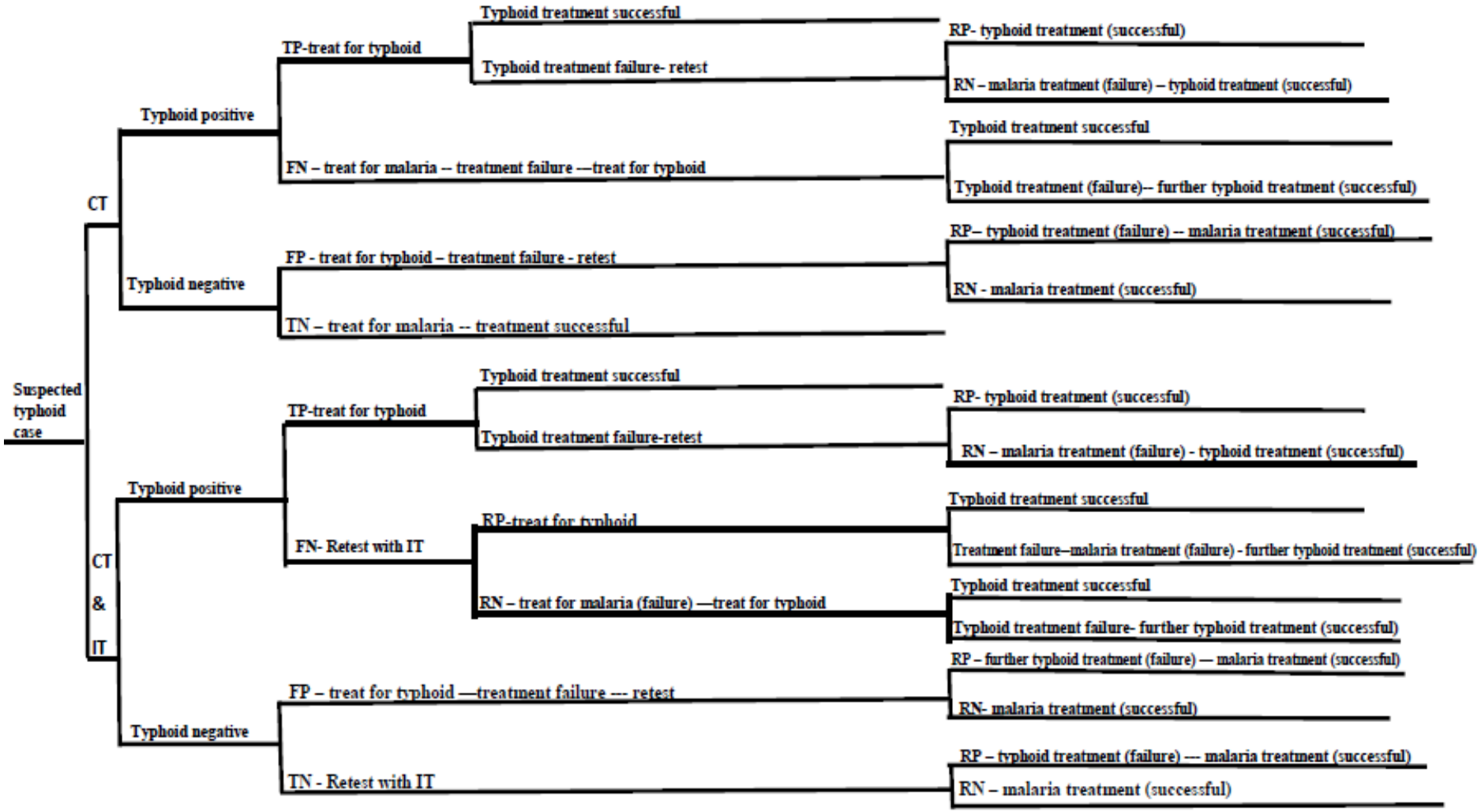
Appendix 5: Models for the economic analysis

Fig A5-1 Model comparing the HT-test vs. Widal test in the triage role



Where CT refers to comparator test (Widal test) Where IT refers to Index test (the HT-test)

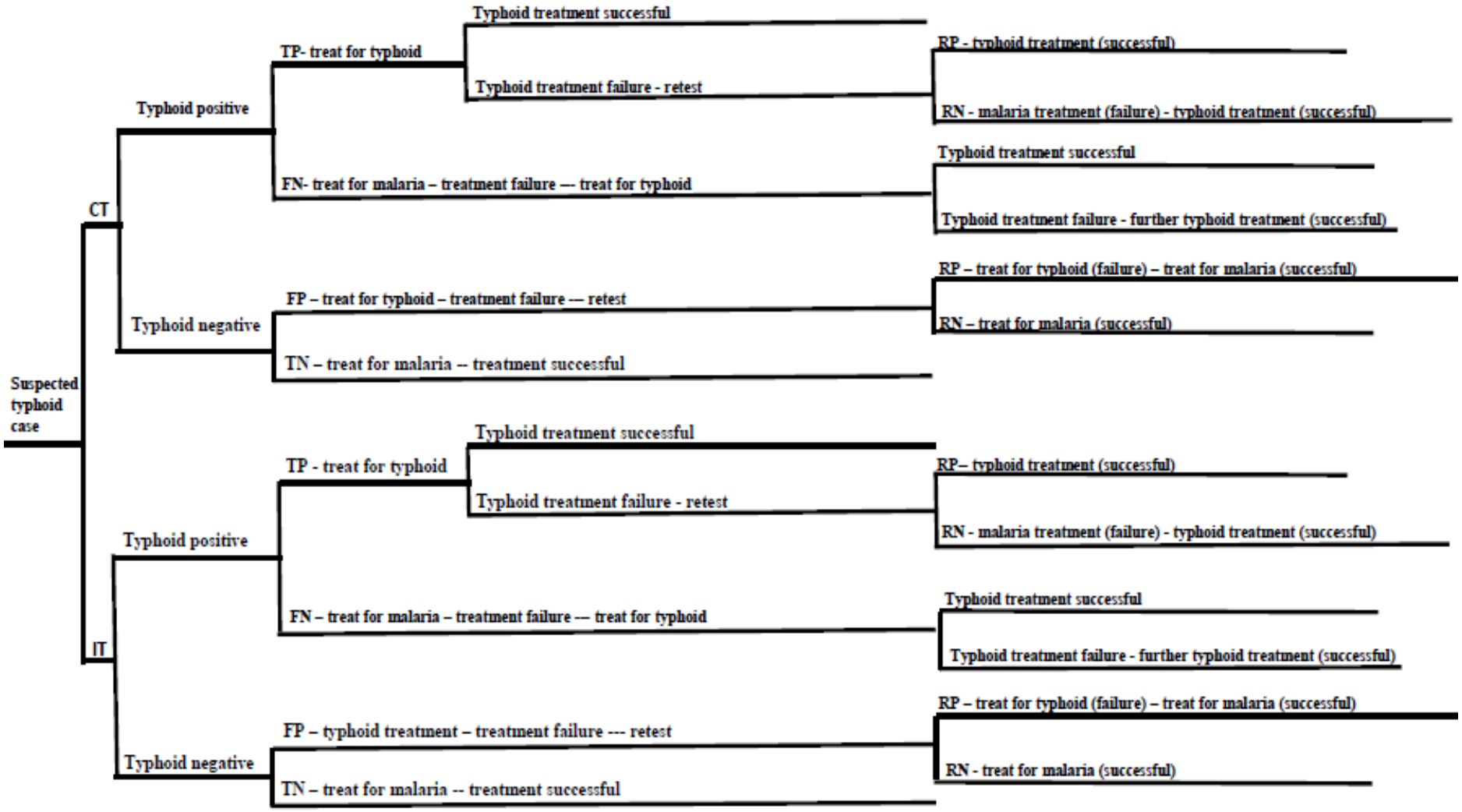
Fig A5-2 Model comparing the HT-test vs. Widal test in the add-on role



Where CT refers to comparator test (Widal test)

Where IT refers to Index tests (the HT-tests)

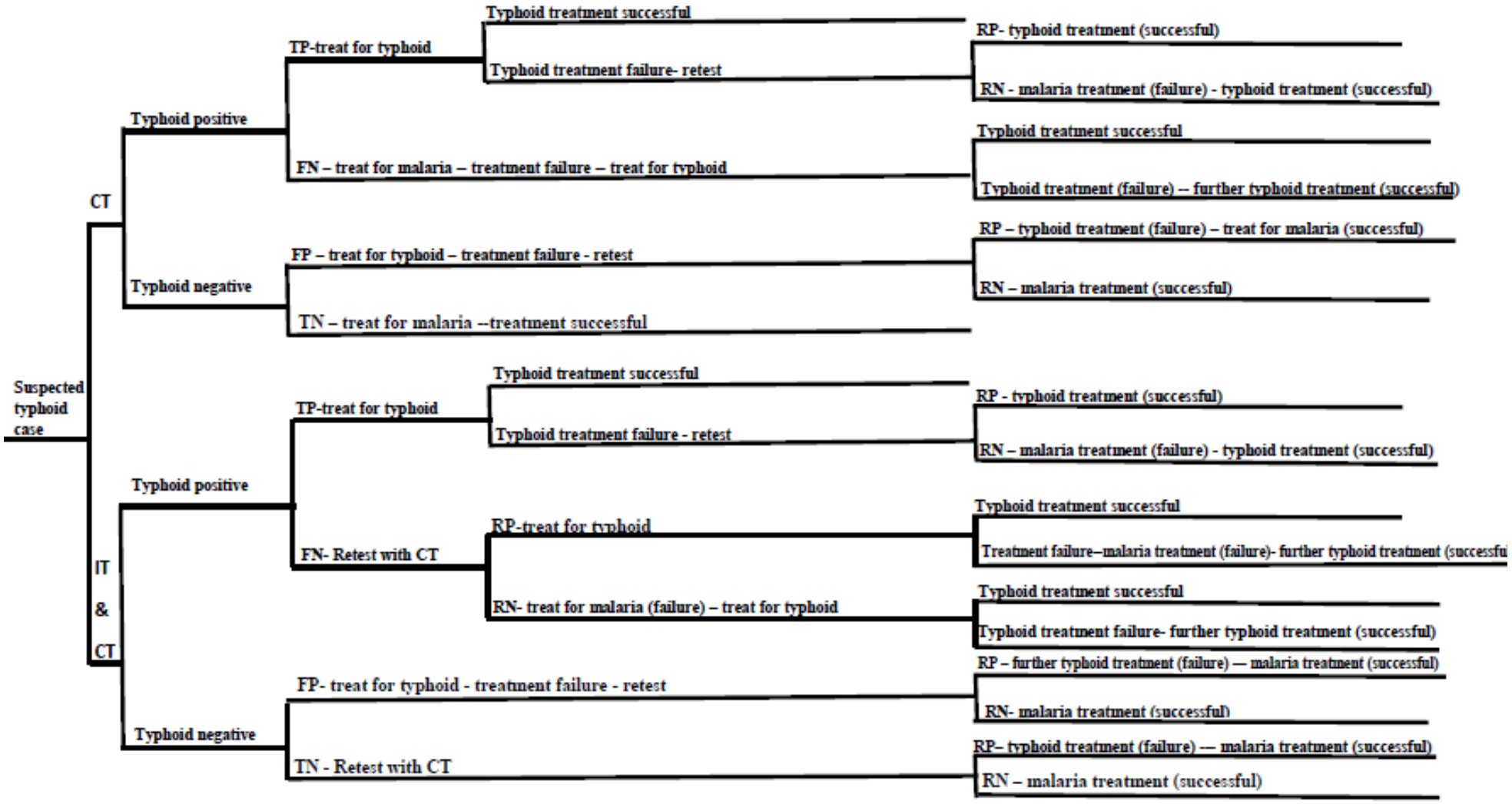
Fig A5-3 Model comparing the HT-test vs. blood culture in the replacement role



Where CT refers to comparator test (blood culture)

Where IT refers to Index test (the HT-test)

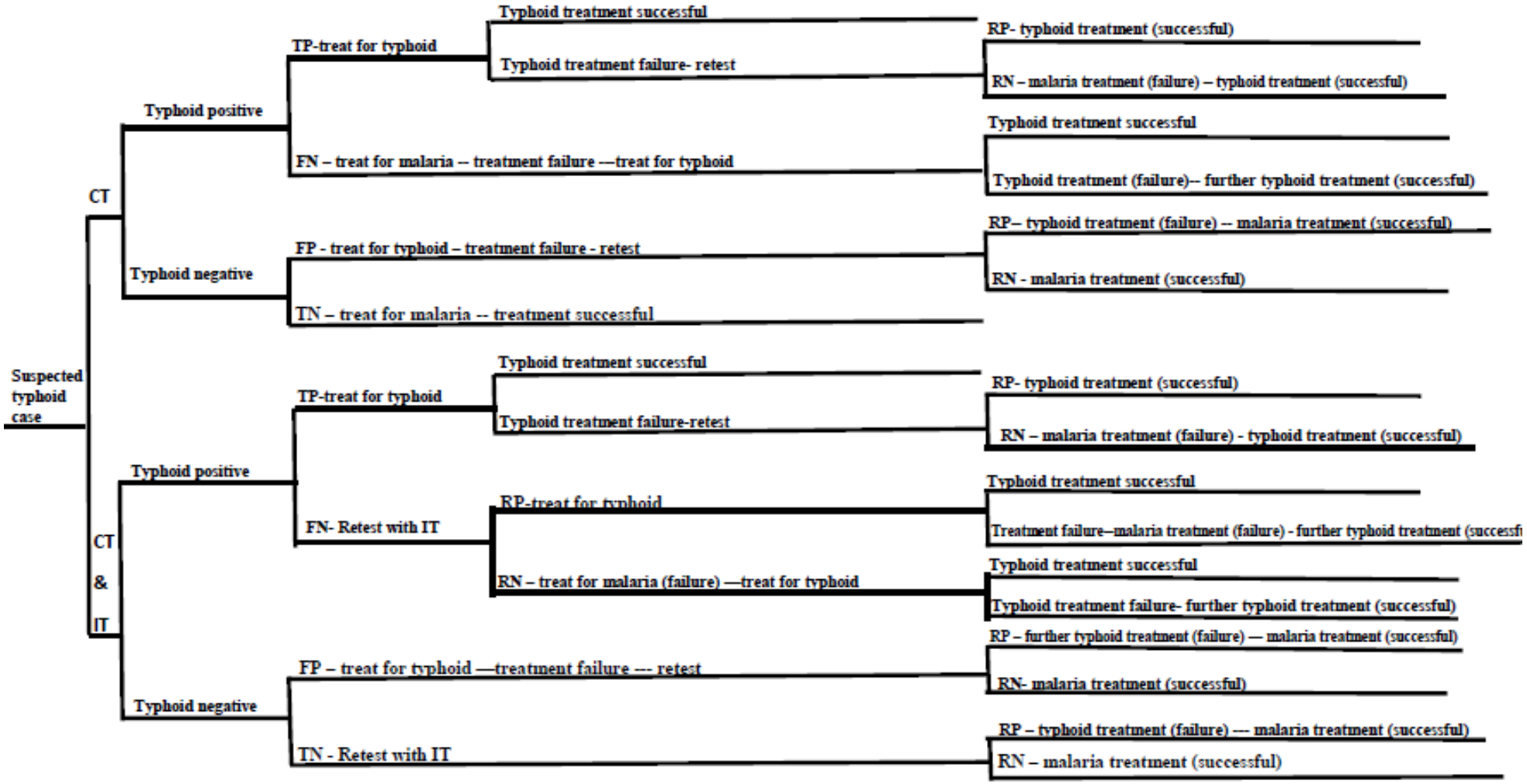
Fig A5-4 Model comparing the HT-test vs. blood culture in the triage role



Where CT refers to comparator test (blood culture)

Where IT refers to Index test (the HT-test)

Fig A5-5 Model comparing the HT-test vs. blood culture in the add-on role



Where CT refers to comparator test (blood culture)

Where IT refers to Index test (the HT-test)

Appendix 6: Headroom analysis (the HT-test vs. Widal test)

Fig A6-1 Headroom of the HT-test vs. Widal test at a WTP of $951/QALY (Triage role)

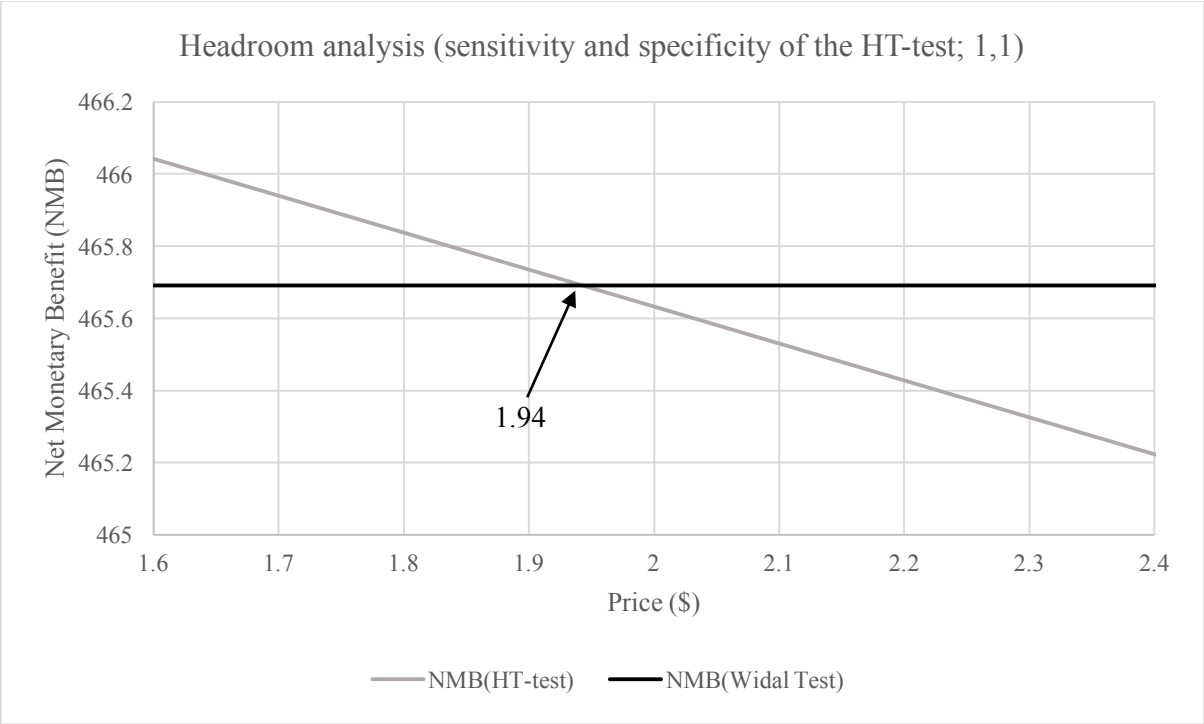
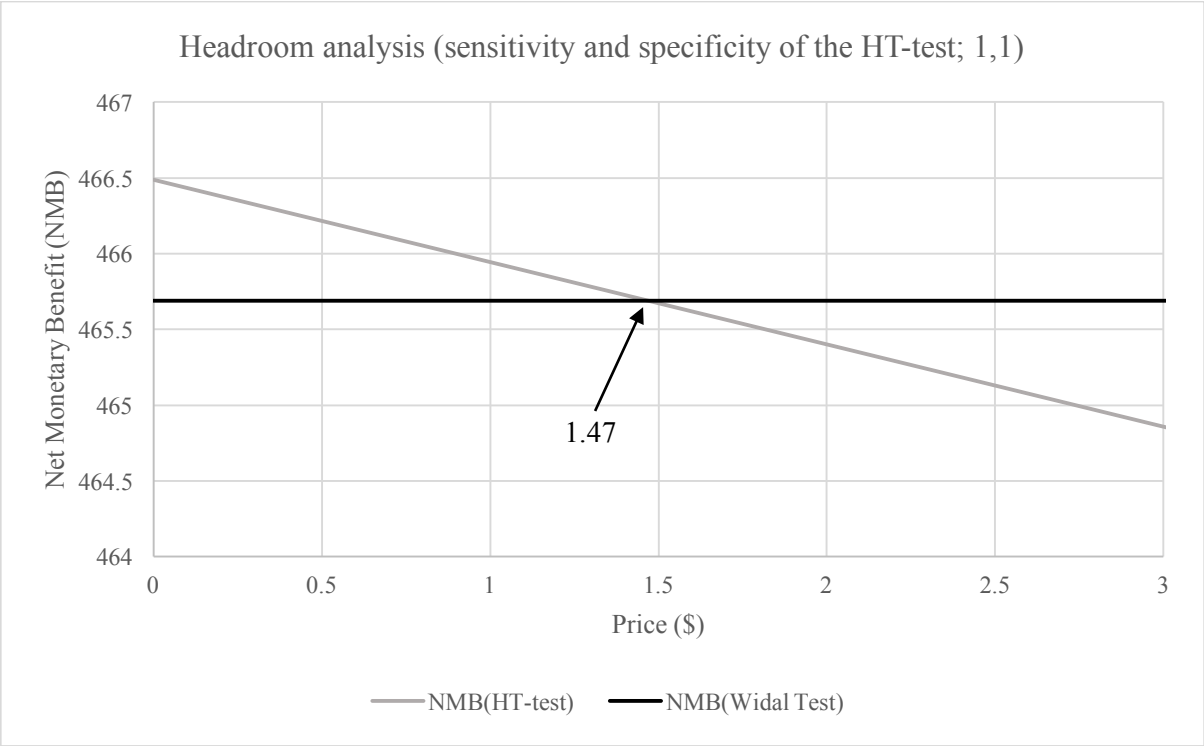


Fig A6-2 Headroom of the HT-test vs. Widal test at a WTP of $951/QALY (Add-on role)



Appendix 7: Incremental effectiveness (QALYs) and headroom (the HT-test vs Widal)

Table A7-1 Incremental effectiveness (QALYs) at each sensitivity and specificity pair for the HT-test vs. Widal test (Triage Role)

Sensitivity \ Specificity	100	90	80	70	60	50	40	30	20	10	0
100	0.0003	0.0001	-0.0002	-0.0005	-0.0009	-0.0013	-0.0018	-0.0024	-0.0030	-0.0036	
90	0.0003	0.0001	-0.0002	-0.0005	-0.0009	-0.0014	-0.0019	-0.0024	-0.0030		
80	0.0003	0.0001	-0.0002	-0.0005	-0.0009	-0.0014	-0.0019	-0.0024			
70	0.0003	0.0001	-0.0002	-0.0006	-0.0010	-0.0014	-0.0019				
60	0.0003	0	-0.0002	-0.0006	-0.0010	-0.0014					
50	0.0002	0	-0.0003	-0.0006	-0.0010						
40	0.0002	0	-0.0003	-0.0006							
30	0.0002	0	-0.0003								
20	0.0002	0									
10	0.0002										
0											

Table A7-2 Maximum price ($) at which the HT-test is still cost-effective at each sensitivity and specificity pair for the HT-test vs. Widal test (Triage role)

Sensitivity \ Specificity	100	90	80	70	60	50	40	30	20	10	0
100	**1.94**	**1.5**	*1.01*	*0.48*	NV	NV	NV	NV	NV	NV	
90	**1.78**	**1.35**	*0.87*	*0.34*	NV	NV	NV	NV	NV		
80	**1.63**	**1.2**	*0.73*	*0.20*	NV	NV	NV	NV			
70	**1.47**	**1.06**	*0.58*	*0.07*	NV	NV	NV				
60	**1.32**	**0.91**	*0.44*	NV	NV	NV					
50	**1.17**	**0.77**	*0.31*	NV	NV						
40	**1.02**	**0.62**	*0.17*	NV							
30	**0.88**	**0.48**	*0.03*								
20	**0.73**	**0.34**									
10	**0.59**										
0											

Table A7-3 Incremental effectiveness (QALYS) at each sensitivity and specificity pair for the HT-test vs. Widal test (Add-on role)

Sensitivity	100	90	80	70	60	50	40	30	20	10	0
100	0.0002	0	-0.0002	-0.0004	-0.0006	-0.0008	-0.0010	-0.0013	-0.0015	-0.0017	
90	0.0001	-0.0001	-0.0003	-0.0005	-0.0007	-0.0009	-0.0011	-0.0013	-0.0015		
80	0.0001	-0.0001	-0.0003	-0.0005	-0.0007	-0.0009	-0.0011	-0.0013			
70	0.0001	-0.0001	-0.0003	-0.0005	-0.0007	-0.0009	-0.00109709				
60	0.0001	-0.0001	-0.0003	-0.0005	-0.0007	-0.0009					
50	0.0001	-0.0001	-0.0003	-0.0005	-0.0007						
40	0.0001	-0.0001	-0.0003	-0.0005							
30	0	-0.0002	-0.0004								
20	0	-0.0002									
10	0										
0											

Specificity

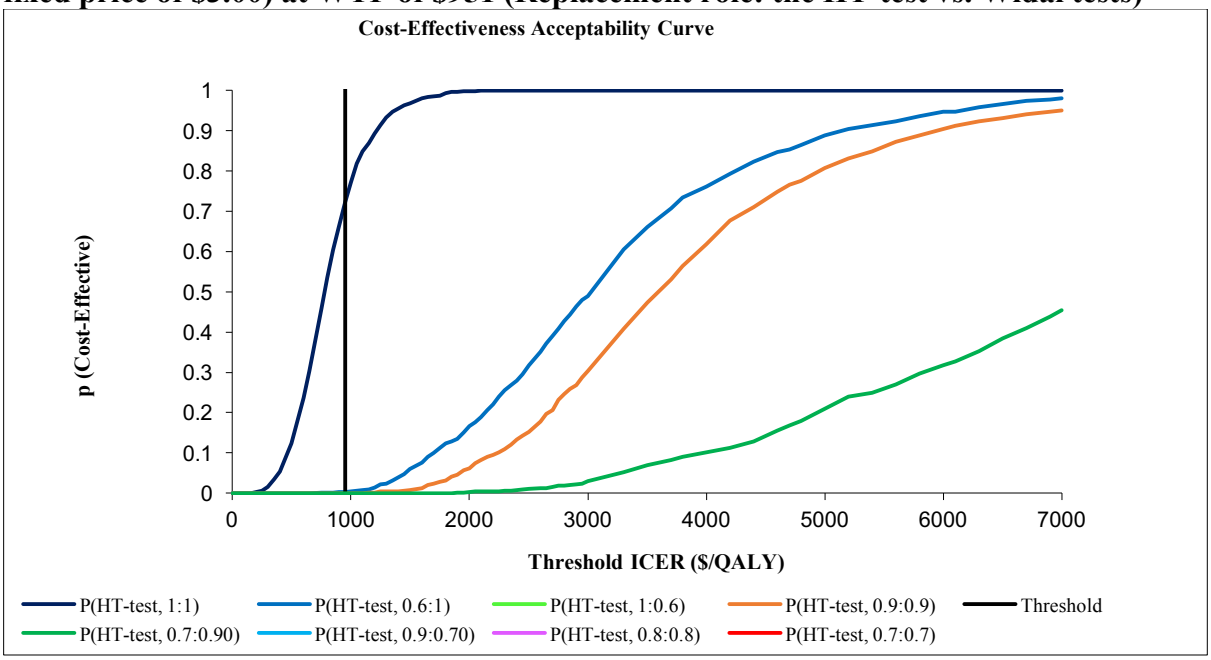
Table A7-4 Maximum price ($) at which HT-test is still cost-effective at each sensitivity and specificity pair for the HT-test vs. Widal test (Add-on role)

Sensitivity	100	90	80	70	60	50	40	30	20	10	0
100	**1.47**	**0.76**	*0.05*	NV	NV	NV	NV	NV	NV	NV	
90	**1.32**	*0.61*	NV	NV	NV	NV	NV	NV	NV		
80	**1.17**	*0.46*	NV	NV	NV	NV	NV	NV			
70	**1.03**	*0.32*	NV	NV	NV	NV	NV				
60	**0.88**	*0.17*	NV	NV	NV	NV					
50	**0.73**	*0.02*	NV	NV	NV						
40	**0.59**	NV	NV	NV							
30	**0.44**	NV	NV								
20	**0.29**	NV									
10	**0.15**										
0											

Specificity

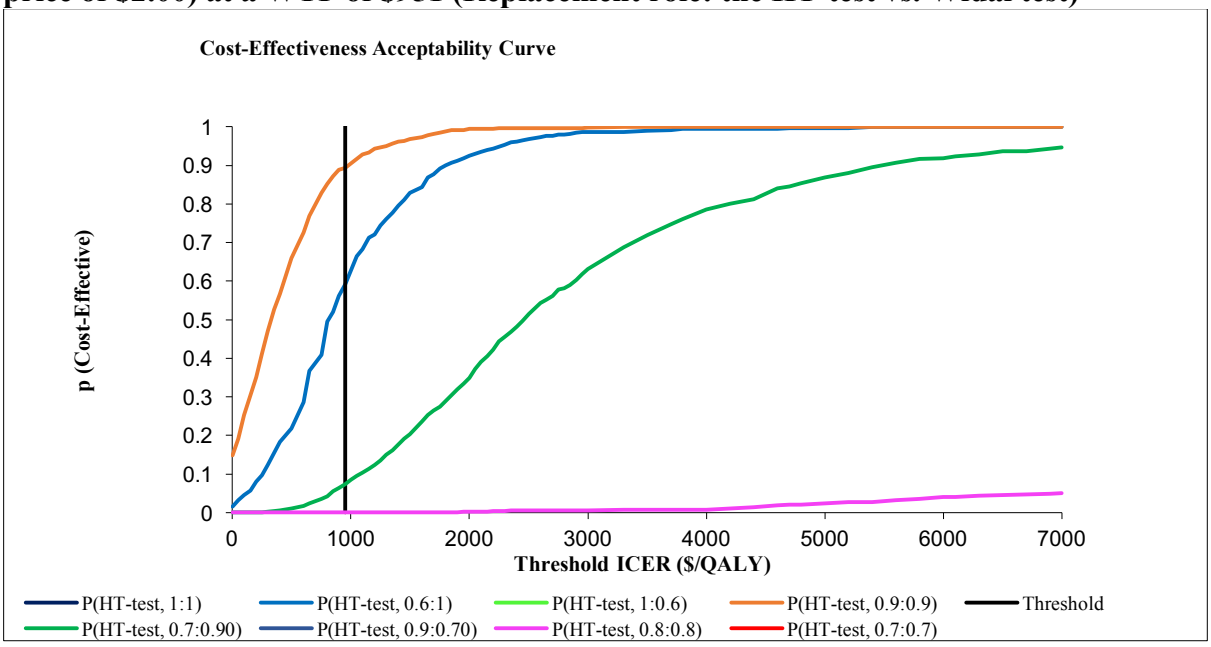
Appendix 8: CEACs of the HT-test vs. Widal test

Fig A8-1 CEAC (with variations in the sensitivity and specificity of the HT-test at a fixed price of $3.00) at WTP of $951 (Replacement role: the HT-test vs. Widal tests)



Curves for the (1:0.6; 0.9:0.7; 0.8:0.8; 0.7:0.7) sensitivity and specificity pairs are not shown because they never go above zero

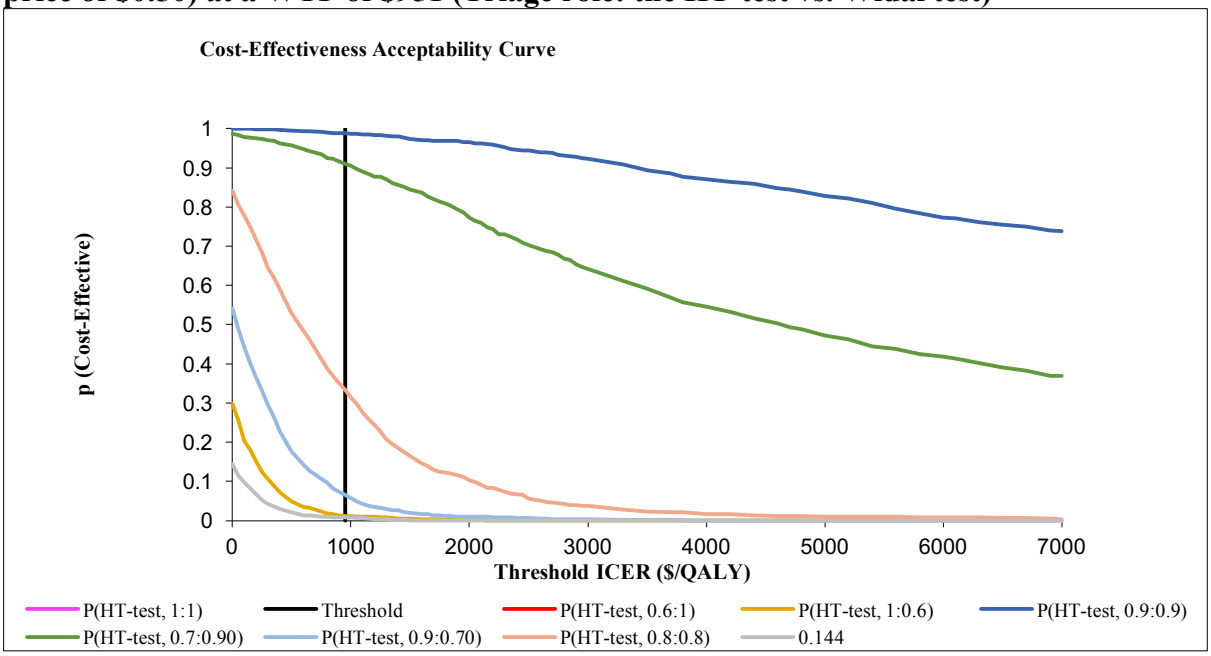
Fig A8-2 CEAC (with variations in the sensitivity and specificity of the HT-test at a fixed price of $2.00) at a WTP of $951 (Replacement role: the HT-test vs. Widal test)



Curves for the (1:0.6; 0.9:0.7; 0.7:0.7) sensitivity and specificity pairs are not shown because they never go above zero

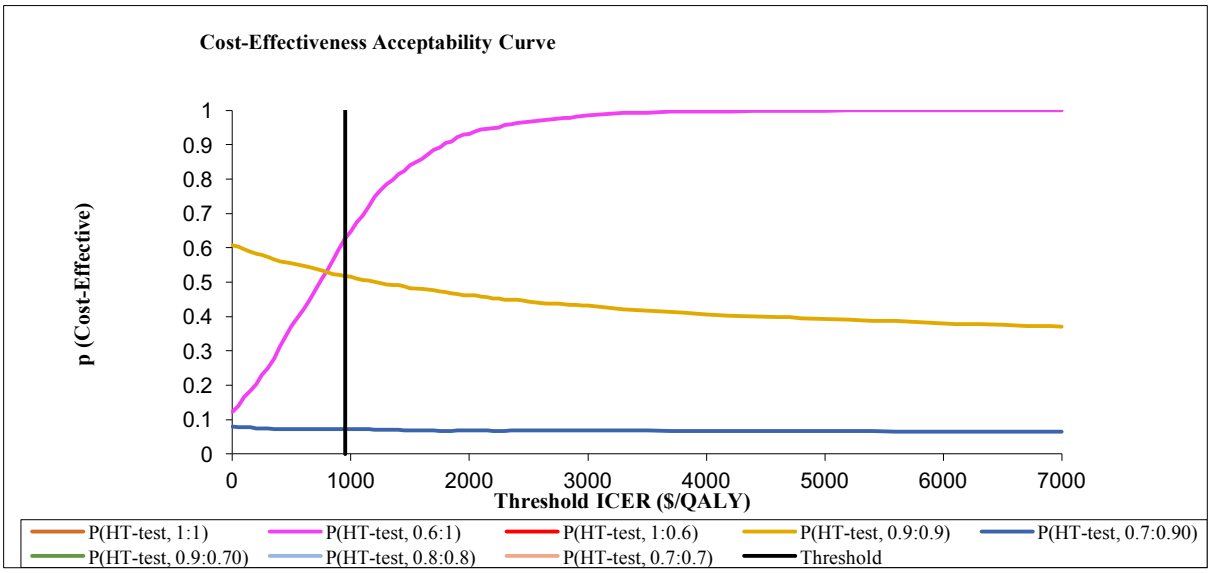
The curve for the (1:1) sensitivity and specificity pair is not shown because it never goes below one

Fig A8-3 CEAC (with variations in the sensitivity and specificity of the HT-test at a fixed price of $0.50) at a WTP of $951 (Triage role: the HT-test vs. Widal test)



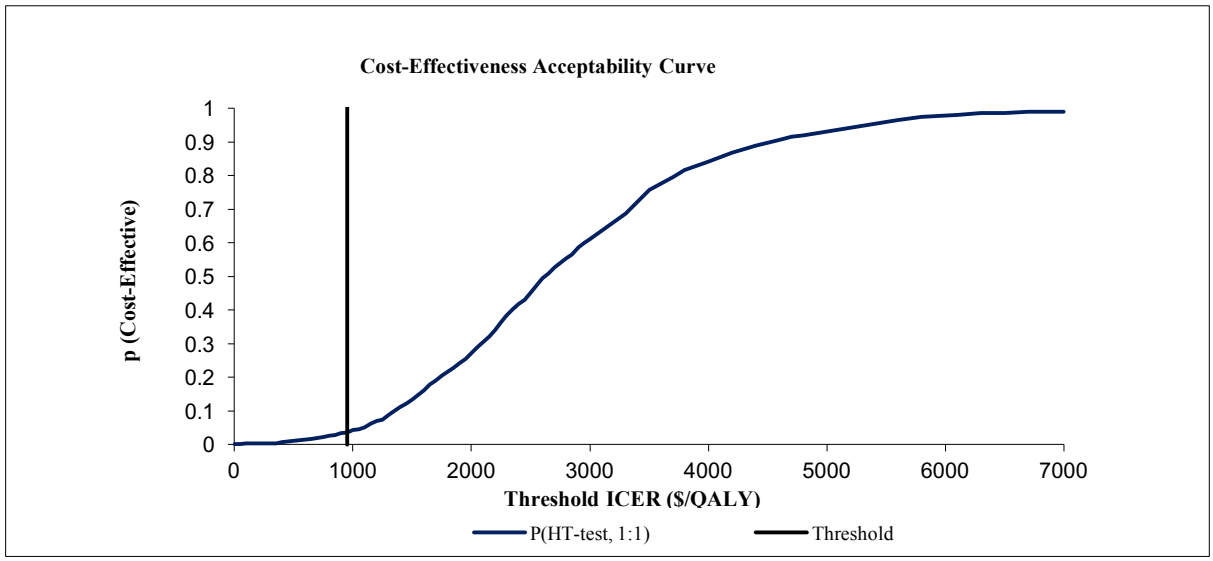
Curves for the (1:1; 0.6:1) sensitivity and specificity pairs are not shown because they never go below one

Fig A8-4 CEAC (with variations in the sensitivity and specificity of the HT-test at a fixed price of $1.00) at a WTP of $951 (Triage role: the HT-test vs. Widal test)



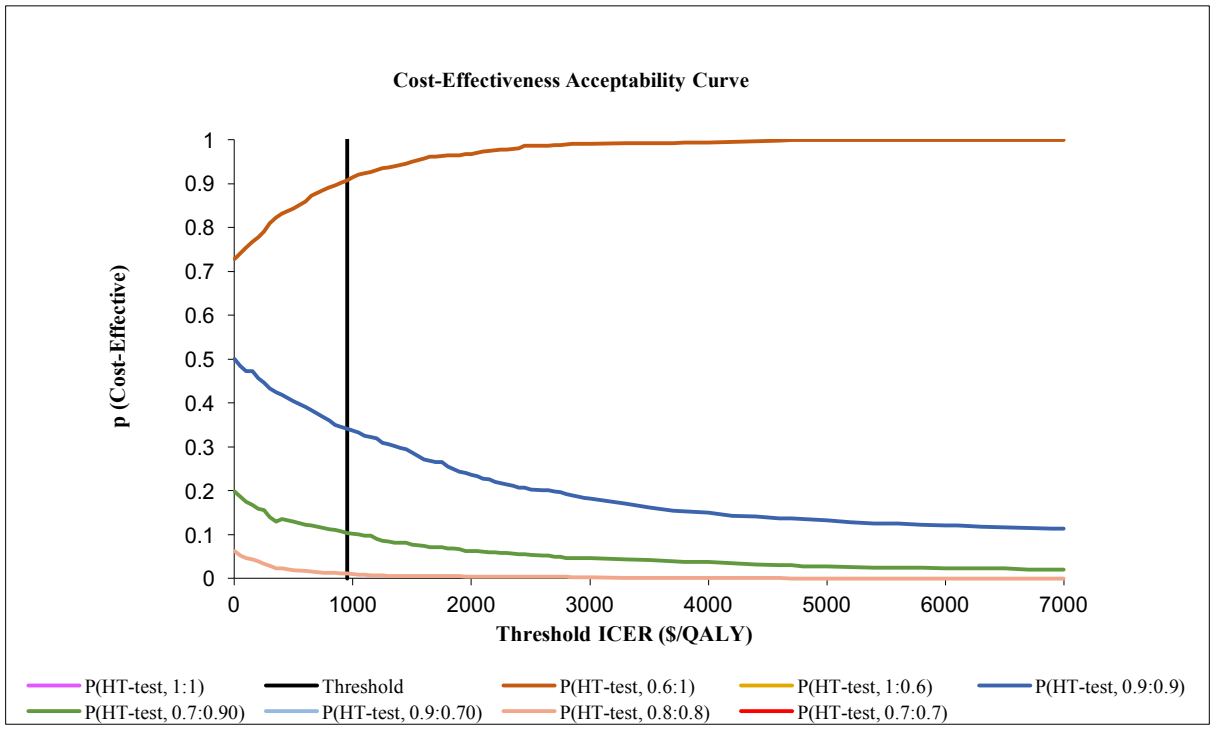
Curves for the (1:0.6; 0.9:0.7; 0.8:0.8; 0.7:0.7) sensitivity and specificity pairs are not shown because they never go above zero
The curve for the (1:1) sensitivity and specificity pair is not shown because it never goes below one

Fig A8-5 CEAC (with variations in the sensitivity and specificity of the HT-test at a fixed price of $2.00) at a WTP of $951 (Triage role: the HT-test vs. Widal test)



Curves for the other sensitivity and specificity pairs are not shown because at a price of $4, only the (1,1) pair has any chance of cost-effectiveness, and even then only above the current threshold.

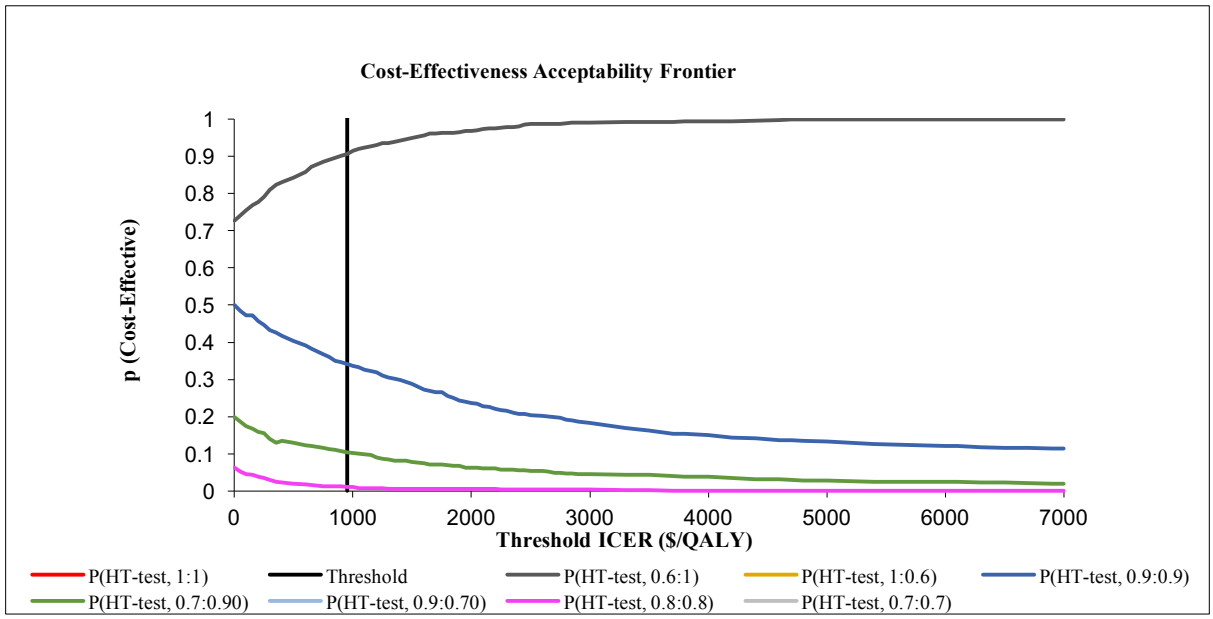
Fig A8-6 CEAC (with variations in the sensitivity and specificity of the HT-test at a fixed price of $0.50) at a WTP of $951 (Add-on role: the HT-test vs. Widal test)



Curves for the (1:0.6; 0.9:0.7; 0.7:0.7) sensitivity and specificity pairs are not shown because they never go above zero

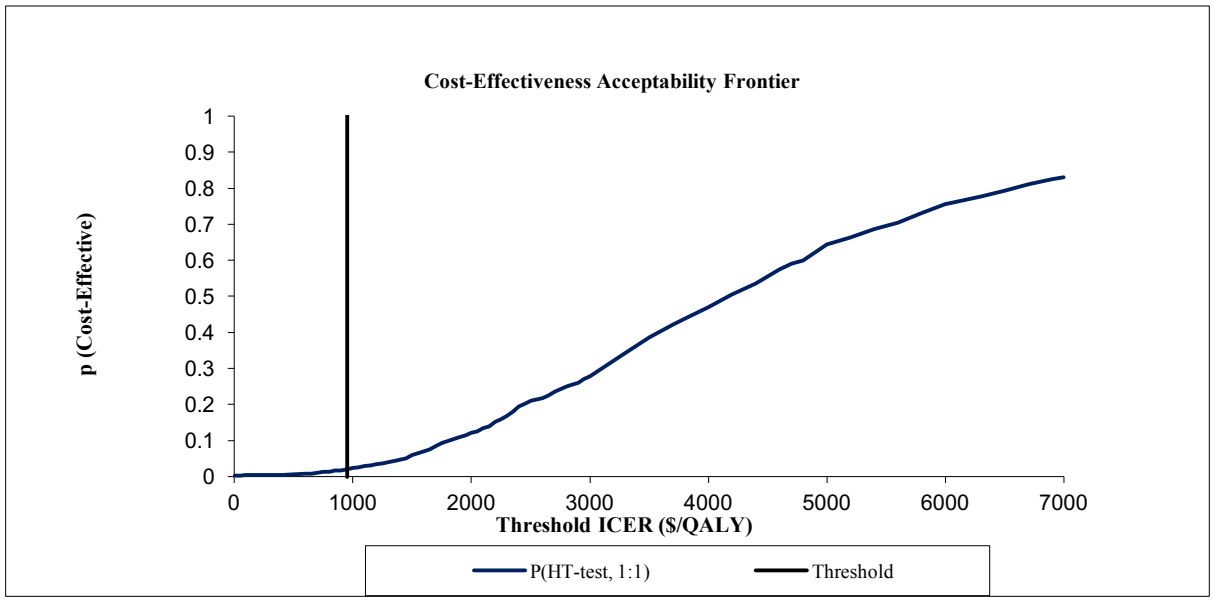
The curve for the (1:1) sensitivity and specificity pair is not shown because it never goes below one

Fig A8-7 CEAC (with variations in the sensitivity and specificity of the HT-test at a fixed price of $1.00) at a WTP of $951 (Add-on role: the HT-test vs. Widal test)



Curves for the (1:0.6; 0.9:0.7; 0.7:0.7) sensitivity and specificity pairs are not shown because they never go above zero: (1:0.6; 0.9:0.7; 0.7:0.7)
The curve for the (1:1) sensitivity and specificity pair is not shown because it never goes below one

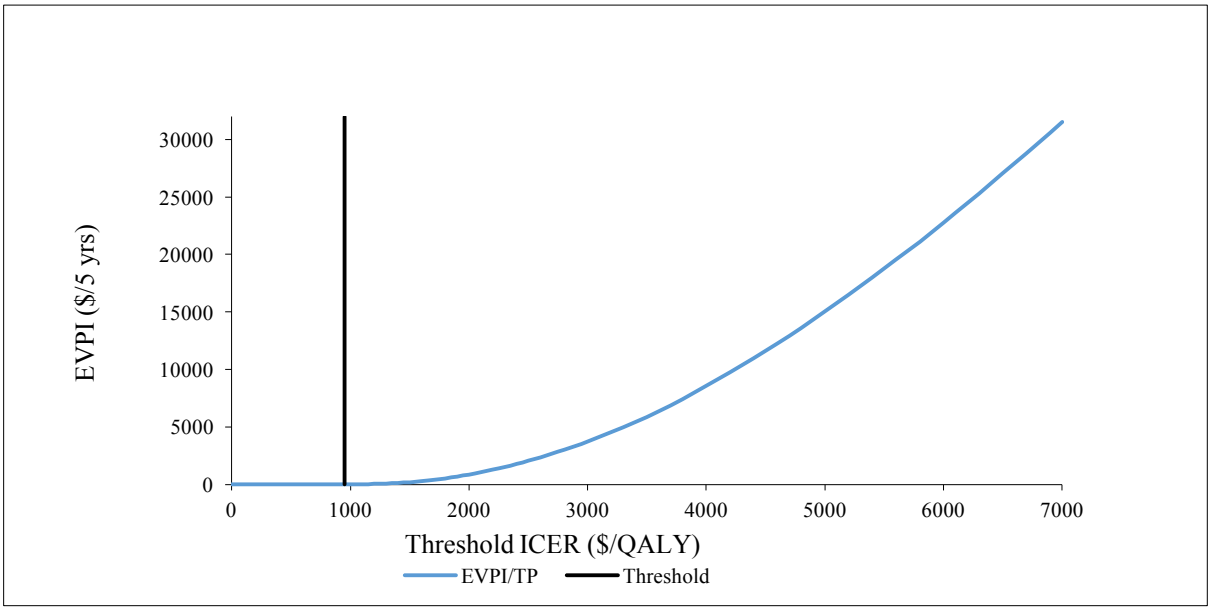
Fig A8-8 CEAC (with variations in the sensitivity and specificity of the HT-test at a fixed price of $2.00) at a WTP of $951 (Add-on role: the HT-test vs. Widal test)



Curves for the other sensitivity and specificity pairs are not shown because at a price of $2, only the (1,1) pair has any chance of cost-effectiveness, and even then only above the current threshold.

Appendix 9: VOI analysis (the HT-test vs. Widal test)

Fig A9-1 VOI analysis of the HT-test vs. Widal test in the replacement role at a price of $2.00



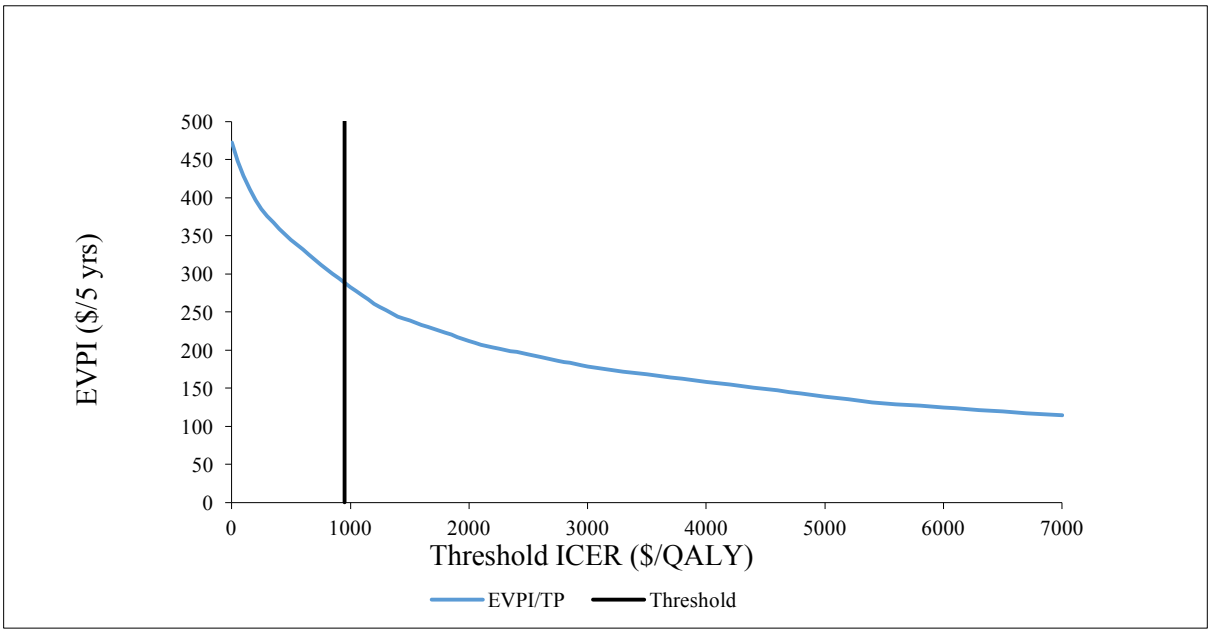
As shown in Fig A9-1, at a WTP threshold of $951/QALY, the overall EVPI for the specified decision relevance time horizon of 5 years is $0.00.

Table A9-1 Single parameter EVPPI when the HT-test is $2.00

Model parameters with uncertainty	Per Person EVPPI ($)	Standard Error	EVPPI for Ghana over 5 years ($)
Mean recovery time for successful typhoid treatment	0.00	0.00	0
Mean recovery time for successful malaria treatment	0.00	0.00	0
Prevalence of typhoid in the patient population presenting	0.00	0.00	0
Probability of first successful typhoid treatment	0.00	0.00	0
Utility when experiencing typhoid symptoms	0.00	0.00	0
Specificity of Widal test	0.00	0.00	0
Sensitivity of Widal test	0.00	0.00	0
Specificity of the HT-test	0.00	0.00	0
Sensitivity of the HT-test	0.00	0.00	0

Table A9-1 shows that none of the parameters contribute to decision uncertainty as the EVPI for the decision relevant time horizon is $0.00.

Fig A9-2 VOI analysis of the HT-test vs. Widal test in the triage role at a price of $1.00



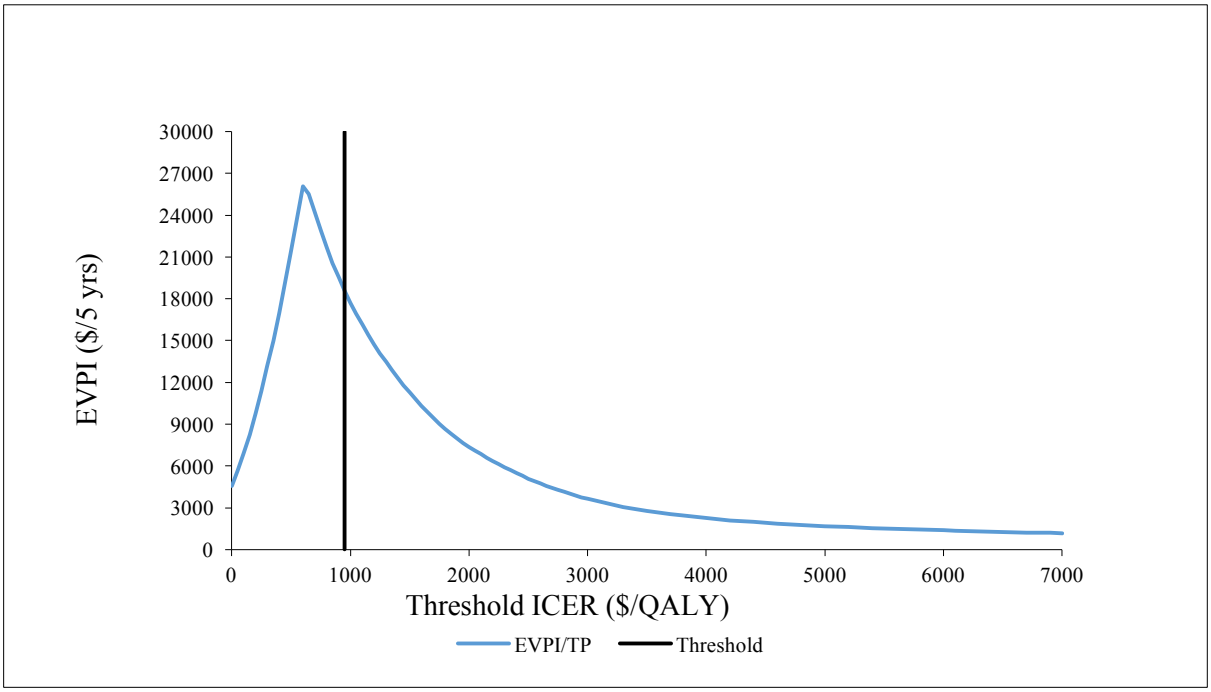
As shown in Fig A9-2, at a WTP threshold of $951/QALY, the overall EVPI for the specified decision relevance time horizon of 5 years is $289.

Table A9-2 Single parameter EVPPI when the HT-test is $1.00

Model parameters with uncertainty	Per Person EVPPI ($)	Standard Error	EVPPI for Ghana over 5 years ($)
Mean recovery time for successful typhoid treatment	0	0	0.00
Mean recovery time for successful malaria treatment	0	0	0.00
Prevalence of typhoid in the patient population presenting	0.000109	0	31
Probability of first successful typhoid treatment	0	0	0.00
Utility when experiencing typhoid symptoms	0	0	0.00
Specificity of Widal test	0	0	0.00
Sensitivity of Widal test	0	0	0.00
Specificity of the HT-test	0	0	0.00
Sensitivity of the HT-test	0	0	0.00

Table A9-2 shows that prevalence of typhoid in the patient population presenting is the parameter contributing to the decision uncertainty in this instance.

Fig A9-3 VOI analysis of the HT-test vs. Widal test in the triage role at a price of $0.5



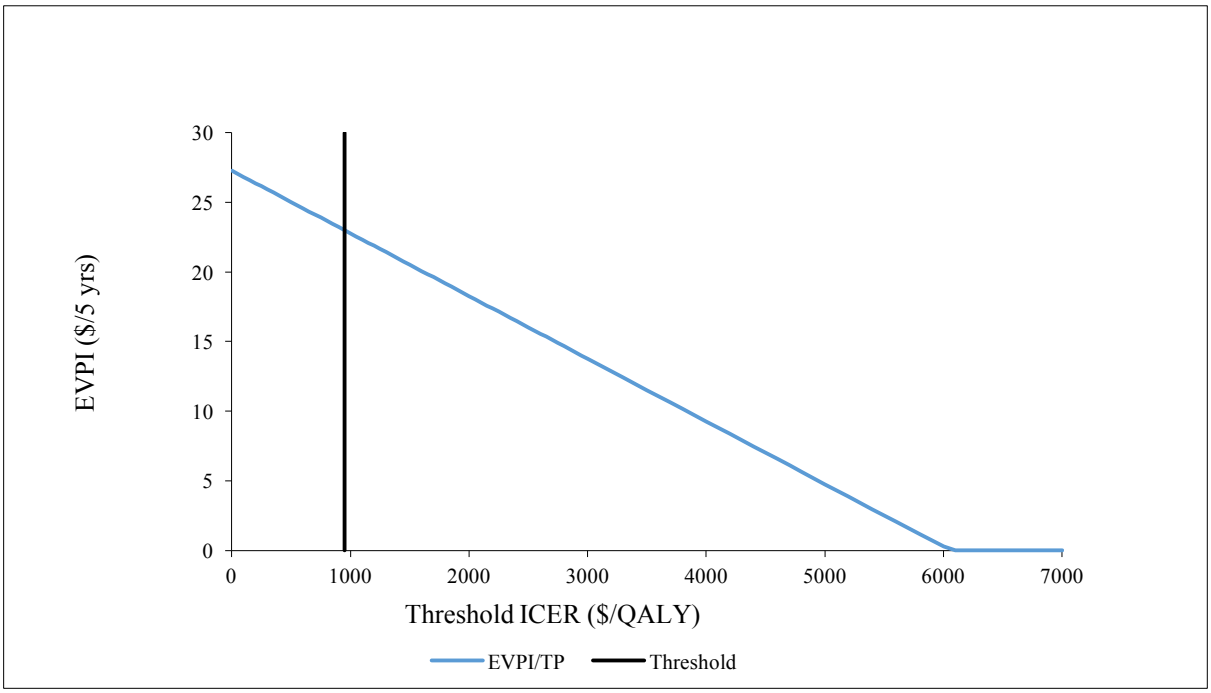
As shown in Fig A9-3, at a WTP threshold of $951/QALY, the overall EVPI for the specified decision relevance time horizon of 5 years is $18585.

Table A9-3 Single parameter EVPPI when the HT-test is $0.5

Model parameters with uncertainty	Per Person EVPPI ($)	Standard Error	EVPPI for Ghana over 5 years ($)
Mean recovery time for successful typhoid treatment	0.00	0	0.000
Mean recovery time for successful malaria treatment	0.00	0	0.000
Prevalence of typhoid in the patient population presenting	0.05	0	14019
Probability of first successful typhoid treatment	0.00	0	0.000
Utility when experiencing typhoid symptoms	0.000024	0	7
Specificity of Widal test	0.000040	0	11
Sensitivity of Widal test	0.000265	0	74
Specificity of the HT-test	0.010029	0	2812
Sensitivity of the HT-test	0.00	0	0.000

Table A9-3 shows that the prevalence of typhoid in the patient population presenting, utility when experiencing typhoid symptoms, specificity and sensitivity of Widal test and the specificity of the HT-test are the parameters contributing to most of the decision uncertainty.

Fig A9-4 VOI analysis of the HT-test vs. Widal test in the add-on role at a price of $1.00



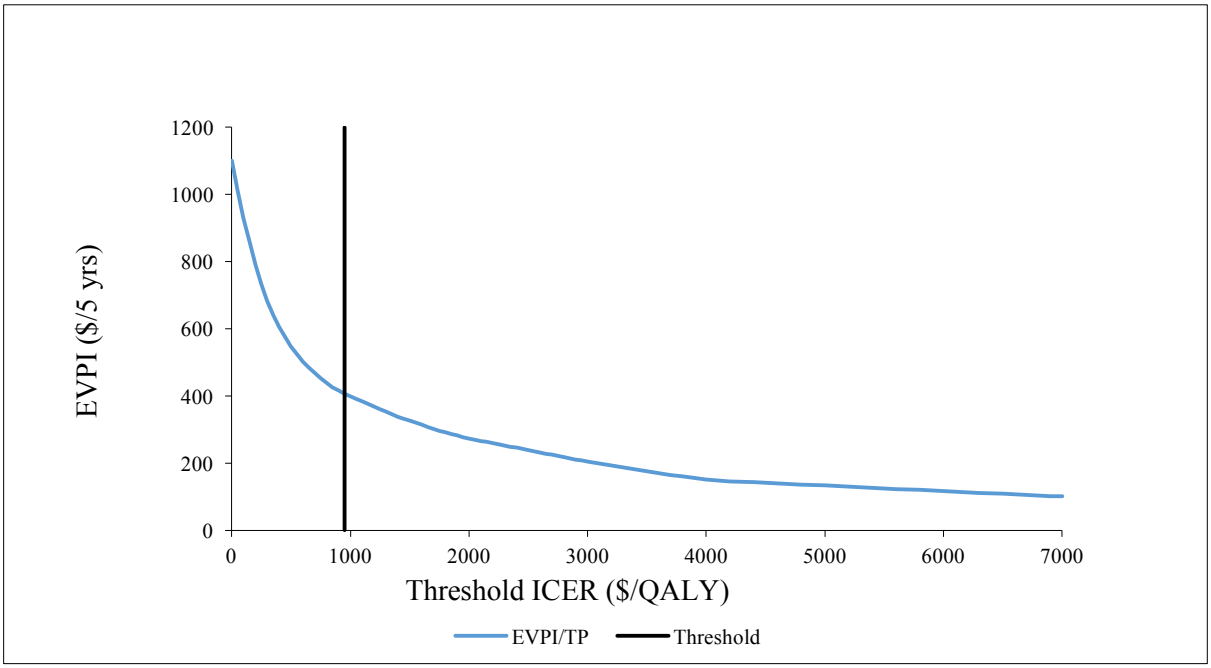
As shown in Fig A9-4, at a WTP threshold of $951/QALY, the overall EVPI for the specified decision relevance time horizon of 5 years is $23.

Table A9-4 Single parameter EVPPI when the HT-test is $1.00

Model parameters with uncertainty	Per Person EVPPI ($)	Standard Error	EVPPI for Ghana over 5 years ($)
Mean recovery time for successful typhoid treatment	0	0	0
Mean recovery time for successful malaria treatment	0	0	0
Prevalence of typhoid in the patient population presenting	0	0	0
Probability of first successful typhoid treatment	0	0	0
Utility when experiencing typhoid symptoms	0	0	0
Specificity of Widal test	0	0	0
Sensitivity of Widal test	0	0	0
Specificity of the HT-test	0	0	0
Sensitivity of the HT-test	0	0	0

Table A9-4 shows that none of the parameters contribute to the decision uncertainty.

Fig A9-5 VOI analysis of the HT-test vs. Widal test in the add-on role at a price of $0.5



As shown in Fig A9-5, at a WTP threshold of $951/QALY, the overall EVPI for the specified decision relevance time horizon of 5 years is $406.

Table A9-5 Single parameter EVPPI when the HT-test is $0.5

Model parameters with uncertainty	Per Person EVPPI ($)	Standard Error	EVPPI for Ghana over 5 years ($)
Mean recovery time for successful typhoid treatment	0	0	0.00
Mean recovery time for successful malaria treatment	0	0	0.00
Prevalence of typhoid in the patient population presenting	0.000307	0	86
Probability of first successful typhoid treatment	0	0	0.00
Utility when experiencing typhoid symptoms	0	0	0.00
Specificity of Widal test	0	0	0.00
Sensitivity of Widal test	0	0	0.00
Specificity of the HT-test	0	0	0.00
Sensitivity of the HT-test	0	0	0.00

Table A9-5 shows that the prevalence of typhoid in the patient population presenting is the parameter contributing to the decision uncertainty.

Appendix 10: Headroom analysis the HT-test vs. blood culture

Fig A10-1 Headroom of HT-test vs blood culture at a WTP of $951/QALY (**Triage role**)

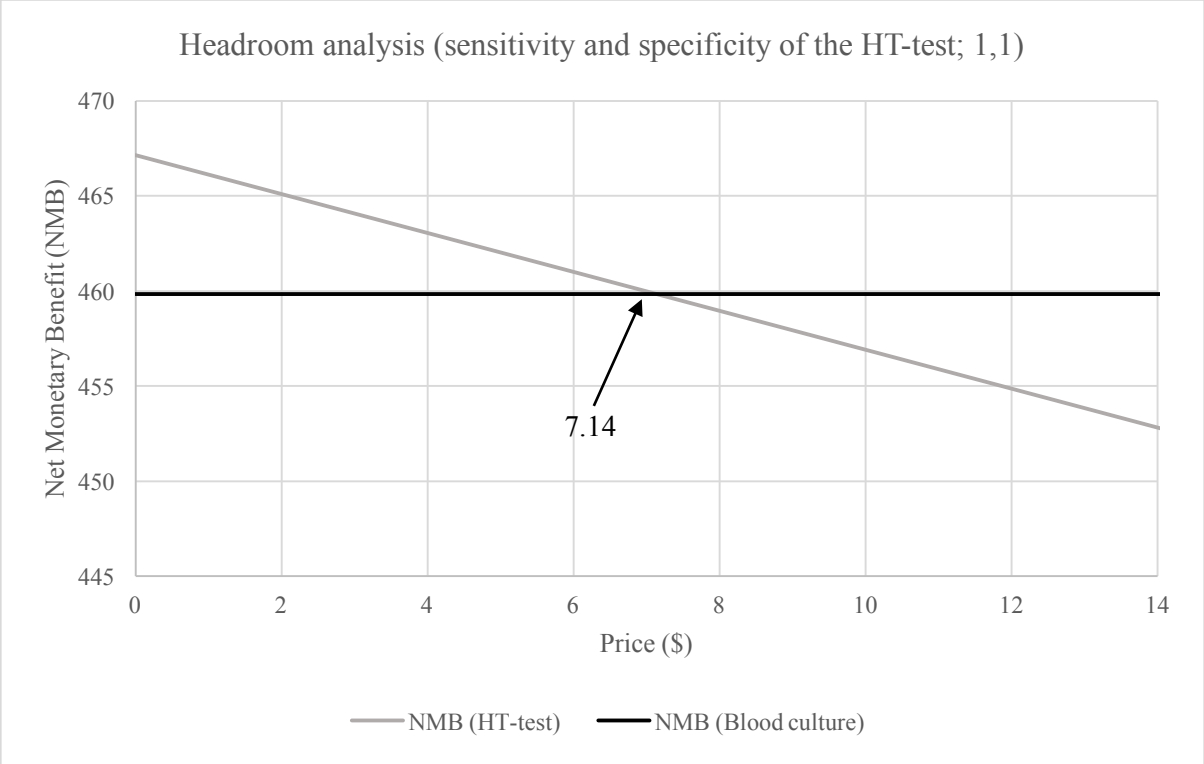
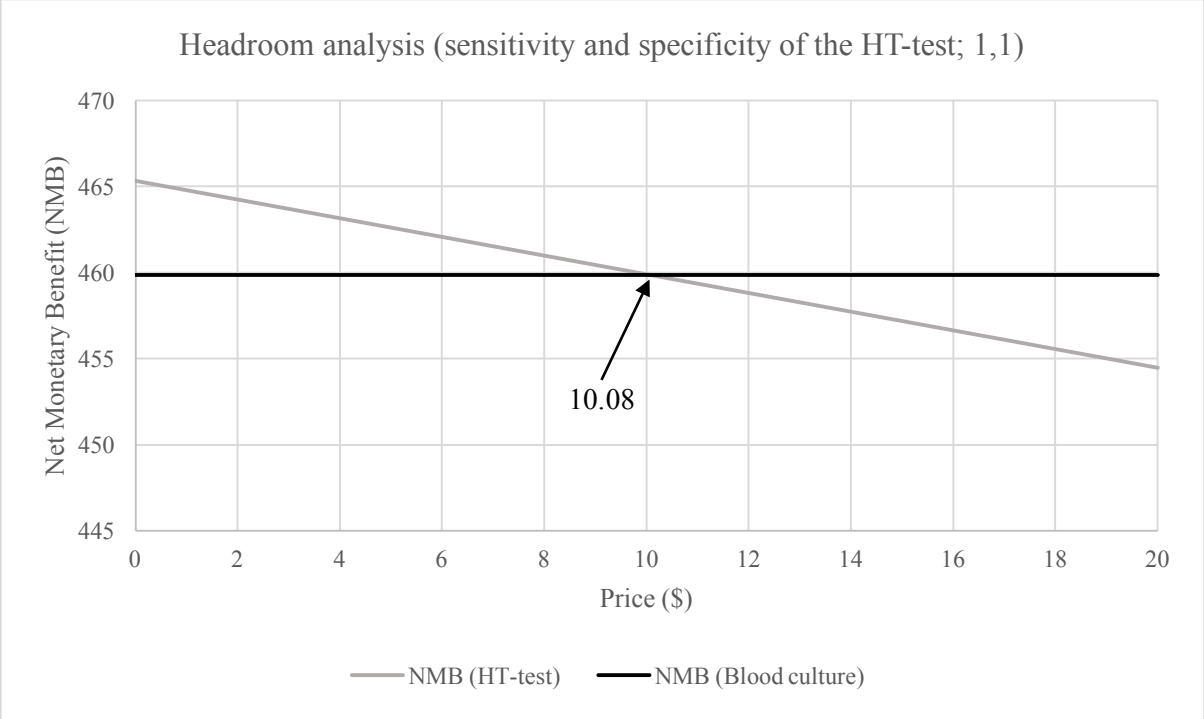


Fig A10-2 Headroom of the HT-test vs. blood culture at a WTP of $951/QALY (**Add-on role**)



Appendix 11: Incremental effectiveness (QALYs) and headroom (HT-test vs blood culture)

Table A11-1 Incremental effectiveness (QALYs) at each sensitivity and specificity pair for HT-test vs blood culture (Triage Role)

Sensitivity	100	90	80	70	60	50	40	30	20	10	0
100	0.0003	0.0001	-0.0002	-0.0005	-0.0009	-0.0013	-0.0018	-0.0024	-0.0030	-0.0036	
90	0.0003	0.0001	-0.0002	-0.0005	-0.0009	-0.0014	-0.0019	-0.0024	-0.0030		
80	0.0003	0.0001	-0.0002	-0.0005	-0.0009	-0.0014	-0.0019	-0.0024			
70	0.0003	0.0001	-0.0002	-0.0006	-0.0010	-0.0014	-0.0019				
60	0.0003	0	-0.0002	-0.0006	-0.0010	-0.0014					
50	0.0002	0	-0.0003	-0.0006	-0.0010						
40	0.0002	0	-0.0003	-0.0006							
30	0.0002	0	-0.0003								
20	0.0002	0									
10	0.0002										
0											

Specificity

Table A11-2 Maximum price at which HT-test is still cost-effective at each sensitivity and specificity pair for HT-test vs blood culture (Triage role)

Sensitivity	100	90	80	70	60	50	40	30	20	10	0
100	**7.14**	**6.5**	*5.82*	*5.12*	*4.4*	*3.65*	*2.89*	*2.1*	*1.31*	*0.5*	
90	**6.94**	**6.31**	*5.64*	*4.95*	*4.23*	*3.49*	*2.73*	*1.95*	*1.16*		
80	**6.75**	**6.13**	*5.47*	*4.78*	*4.07*	*3.33*	*2.58*	*1.8*			
70	**6.56**	**5.94**	*5.29*	*4.61*	*3.9*	*3.17*	*2.42*				
60	**6.37**	**5.76**	*5.12*	*4.44*	*3.74*	*3.02*					
50	**6.19**	**5.58**	*4.95*	*4.28*	*3.58*						
40	**6**	**5.41**	*4.78*	*4.11*							
30	**5.82**	**5.23**	*4.61*								
20	**5.64**	**5.06**									
10	**5.46**										
0											

Specificity

Table A11-3 Incremental effectiveness (QALYs) at each sensitivity and specificity pair for HT-test vs blood culture (Add-on role)

Sensitivity \ Specificity	100	90	80	70	60	50	40	30	20	10	0
100	0.0002	0	-0.0002	-0.0004	-0.0006	-0.0008	-0.0010	-0.0013	-0.0015	-0.0017	
90	0.0001	-0.0001	-0.0003	-0.0005	-0.0007	-0.0009	-0.0011	-0.0013	-0.0015		
80	0.0001	-0.0001	-0.0003	-0.0005	-0.0007	-0.0009	-0.0011	-0.0013			
70	0.0001	-0.0001	-0.0003	-0.0005	-0.0007	-0.0009	-0.0011				
60	0.0001	-0.0001	-0.0003	-0.0005	-0.0007	-0.0009					
50	0.0001	-0.0001	-0.0003	-0.0005	-0.0007						
40	0.0001	-0.0001	-0.0003	-0.0005							
30	0	-0.0002	-0.0004								
20	0	-0.0002									
10	0										
0											

Table A11-4 Maximum price at which HT-test is still cost-effective at each sensitivity and specificity pair for HT-test vs blood culture (Add-on role)

Sensitivity \ Specificity	100	90	80	70	60	50	40	30	20	10	0
100	**10.08**	**9.37**	*8.66*	*7.95*	*7.24*	*6.53*	*5.82*	*5.11*	*4.40*	*3.69*	
90	**9.93**	*9.22*	*8.51*	*7.8*	*7.09*	*6.38*	*5.67*	*4.97*	*4.26*		
80	**9.79**	*9.08*	*8.37*	*7.66*	*6.95*	*6.24*	*5.53*	*4.82*			
70	**9.64**	*8.93*	*8.22*	*7.51*	*6.8*	*6.09*	*5.38*				
60	**9.49**	*8.78*	*8.07*	*7.36*	*6.65*	*5.94*					
50	**9.35**	*8.64*	*7.93*	*7.22*	*6.51*						
40	**9.2**	*8.49*	*7.78*	*7.07*							
30	**9.05**	*8.34*	*7.63*								
20	**8.91**	*8.2*									
10	**8.76**										
0											

Appendix 12: CEACs of the HT-test vs. blood culture

Fig A12-1 CEAC (with variations in the sensitivity and specificity of the HT-test at a fixed price of $9.00) at a WTP of $951 (Replacement role: the HT-test vs. blood culture)

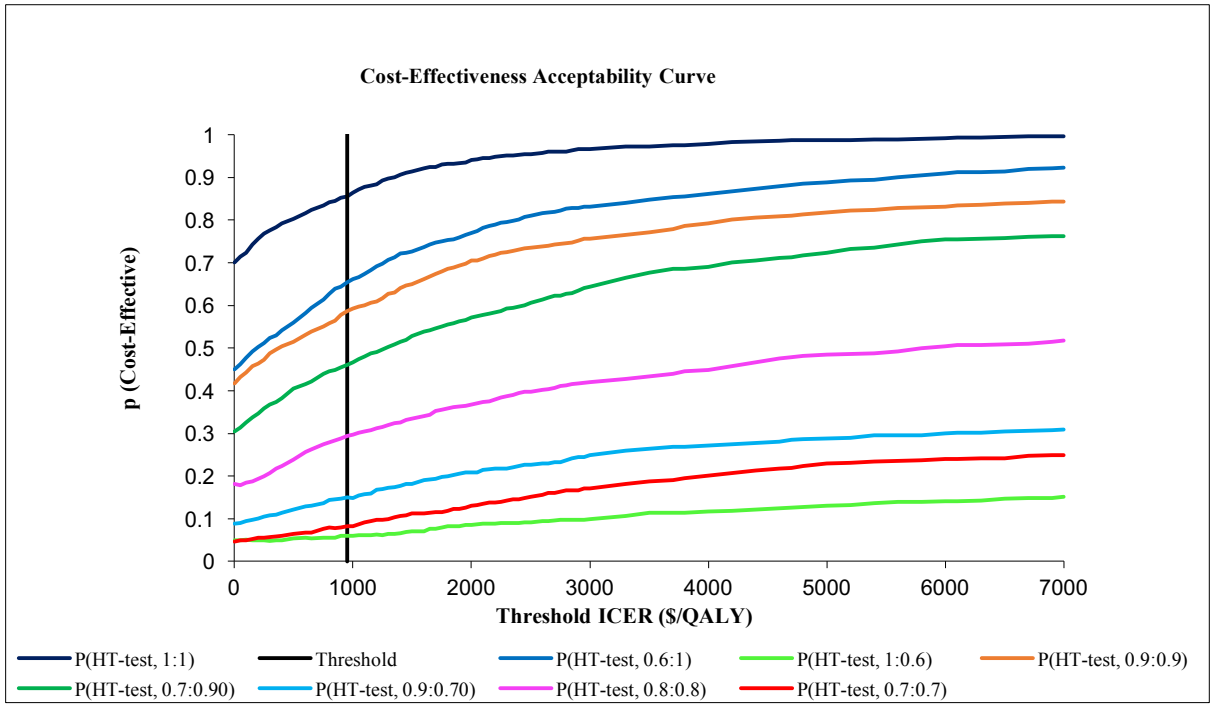


Fig A12-2 CEAC (with variations in the sensitivity and specificity of the HT-test at a fixed price of $8.00) at a WTP of $951 (Replacement role: the HT-test vs. blood culture)

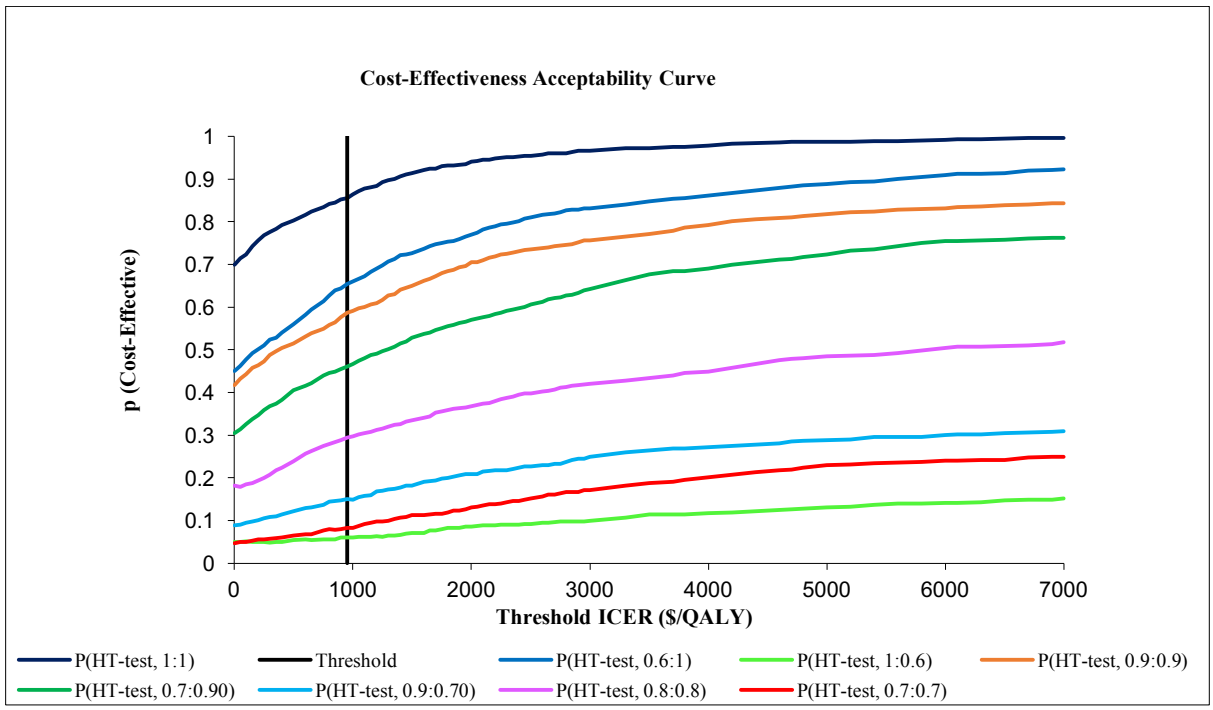


Fig A12-3 CEAC (with variations in the sensitivity and specificity of the HT-test at a fixed price of $7.00) at a WTP of $951 (Replacement role: the HT-test vs. blood culture)

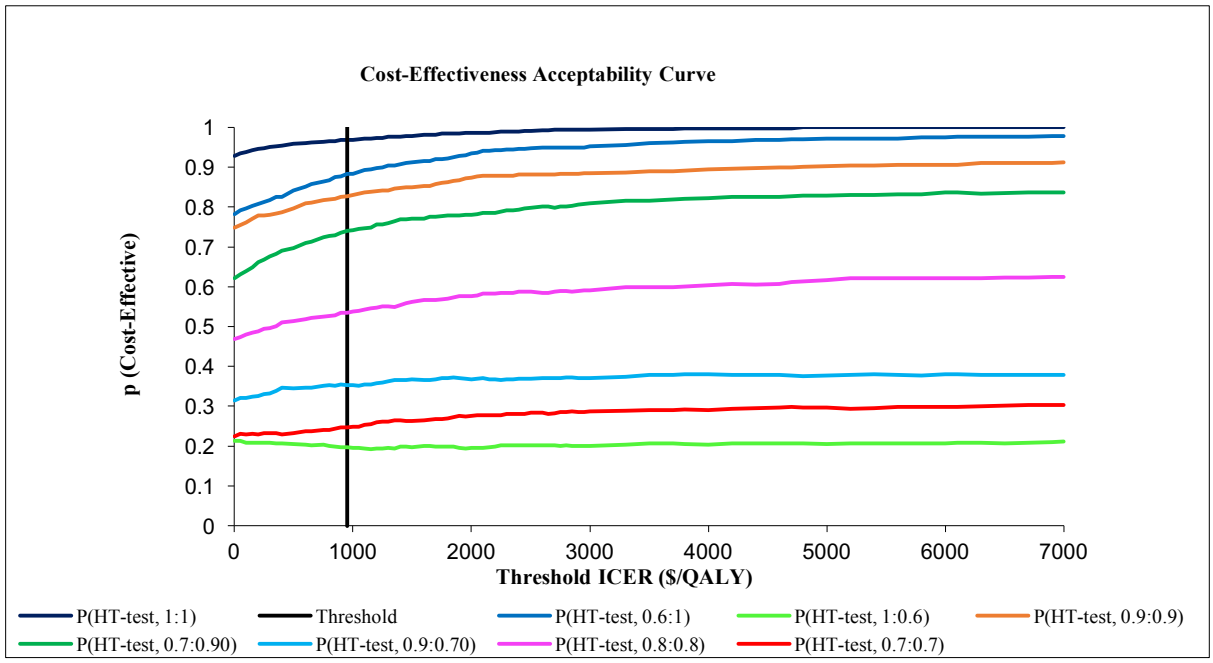


Fig A12-4 CEAC (with variations in the sensitivity and specificity of the HT-test at a fixed price of $5.00) at a WTP of $951 (Triage role: the HT-test vs. blood culture)

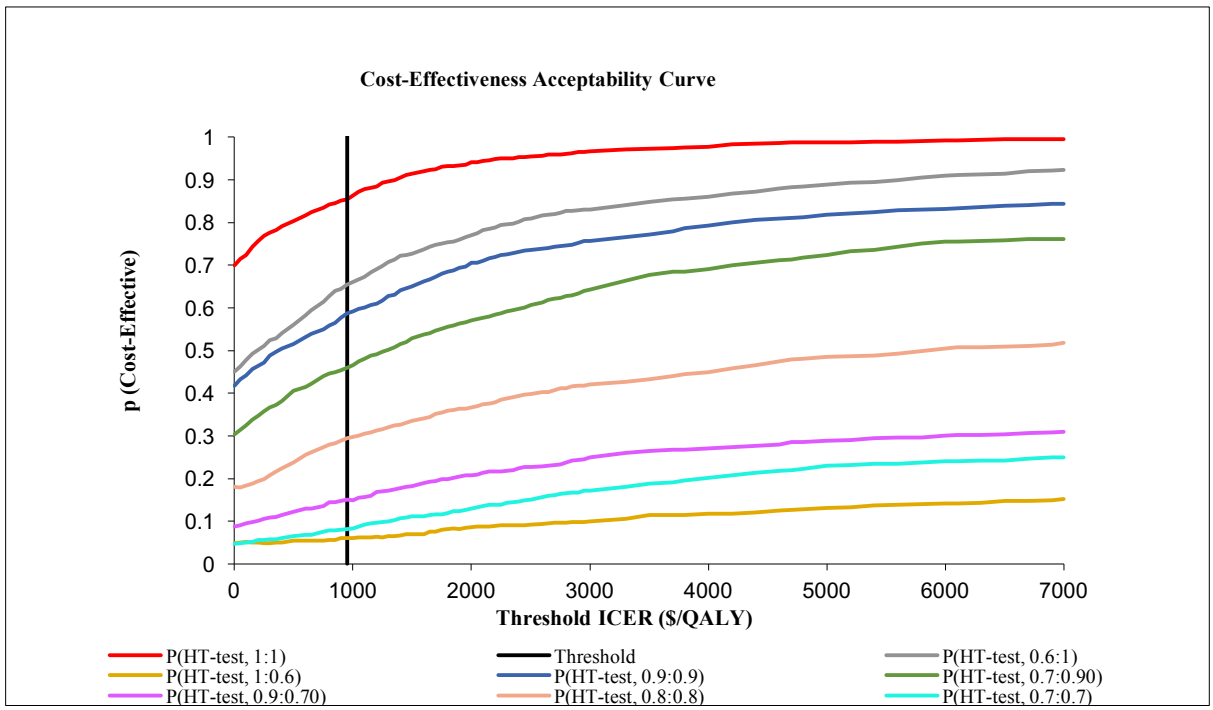


Fig A12-5 CEAC (with variations in the sensitivity and specificity of the HT-test at a fixed price of $6.00) at a WTP of $951 (Triage role: the HT-test vs. blood culture)

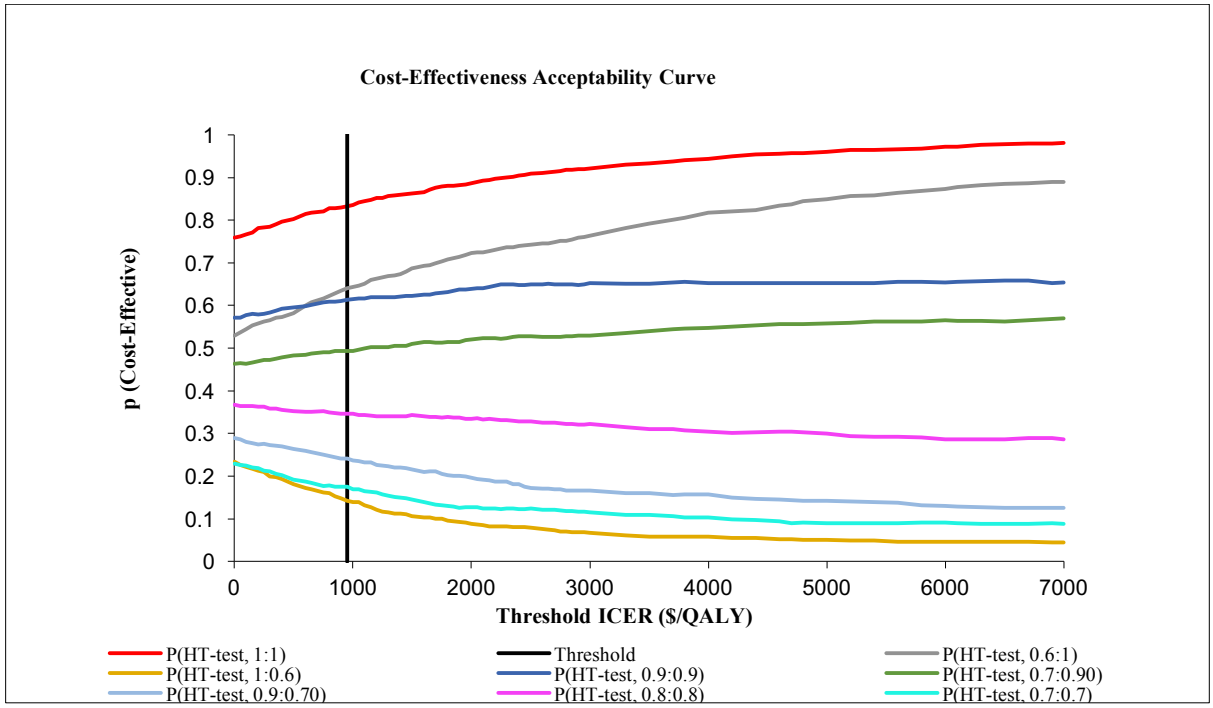


Fig A12-6 CEAC (with variations in the sensitivity and specificity of the HT-test at a fixed price of $7.00) at a WTP of $951 (Triage role: the HT-test vs. blood culture)

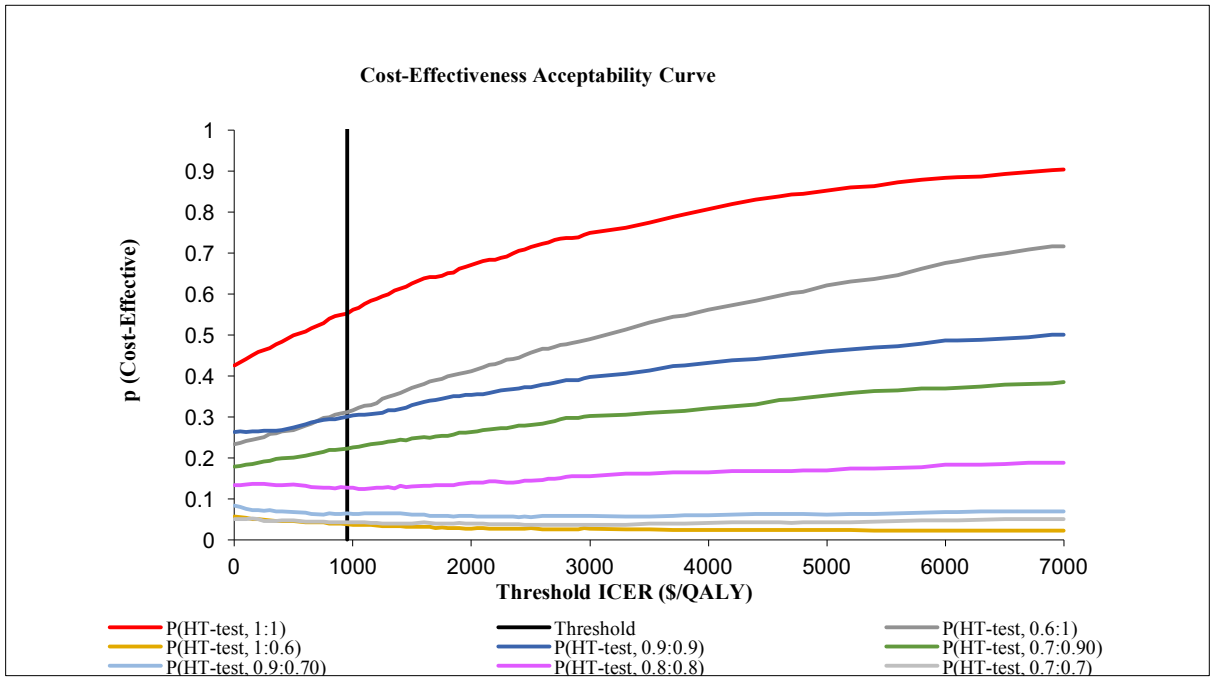
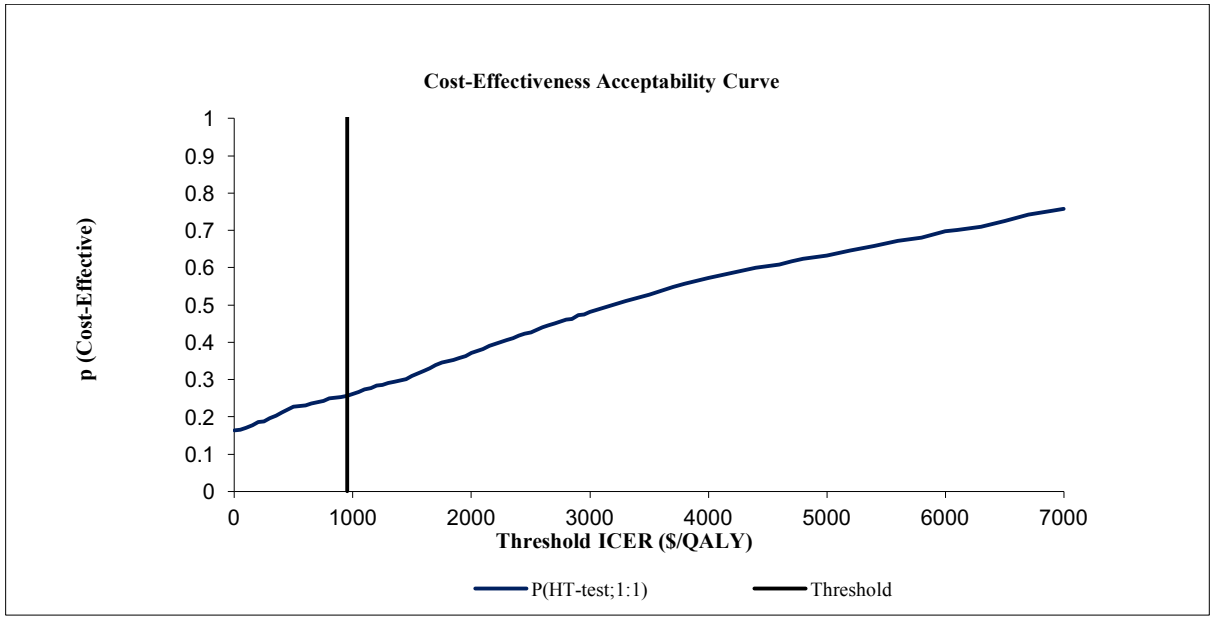


Fig A12-7 CEAC (with variations in the sensitivity and specificity of the HT-test at a fixed price of $7.00) at a WTP of 951 (Triage role: the HT-test vs. blood culture)



Curves for the other sensitivity and specificity pairs are not shown because at a price of $7, only the (1,1) pair has any chance of cost-effectiveness, and even then only above the current threshold.

Fig A12-8 CEAC (with variations in the sensitivity and specificity of the HT-test at a fixed price of $9.00) at a WTP of $951 (Add-on role: the HT-test vs. blood culture)

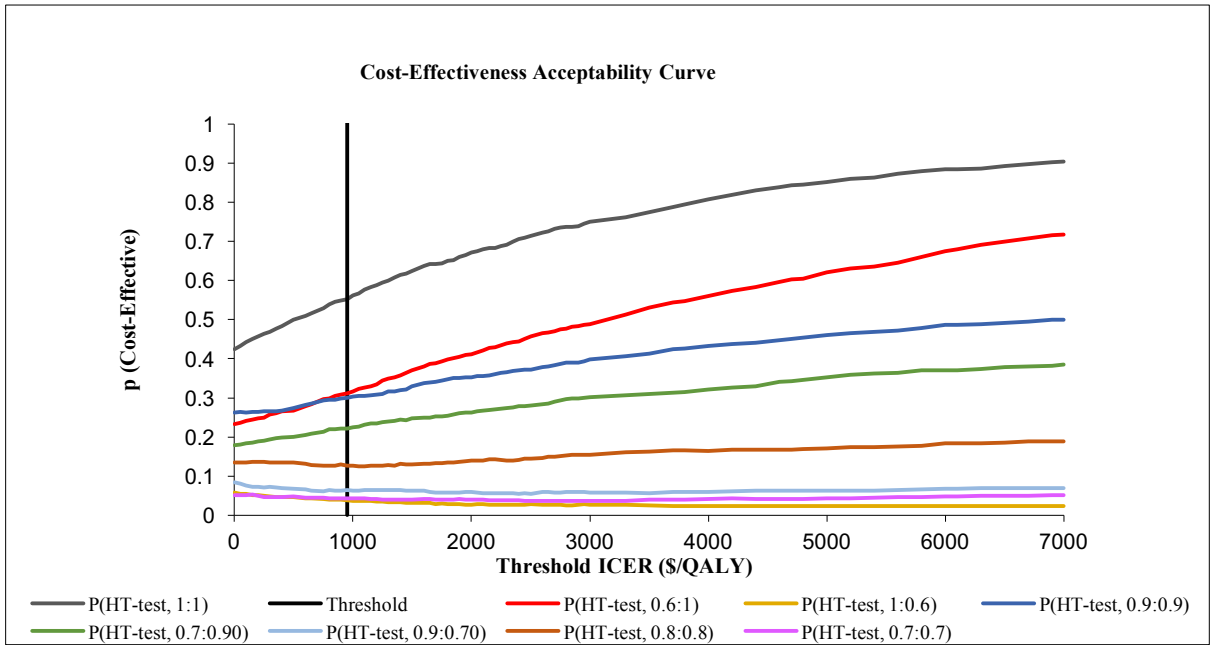


Fig A12-9 CEAC (with variations in the sensitivity and specificity of the HT-test at a fixed price of $10.00) at a WTP of $951 (Add-on role: the HT-test vs. blood culture)

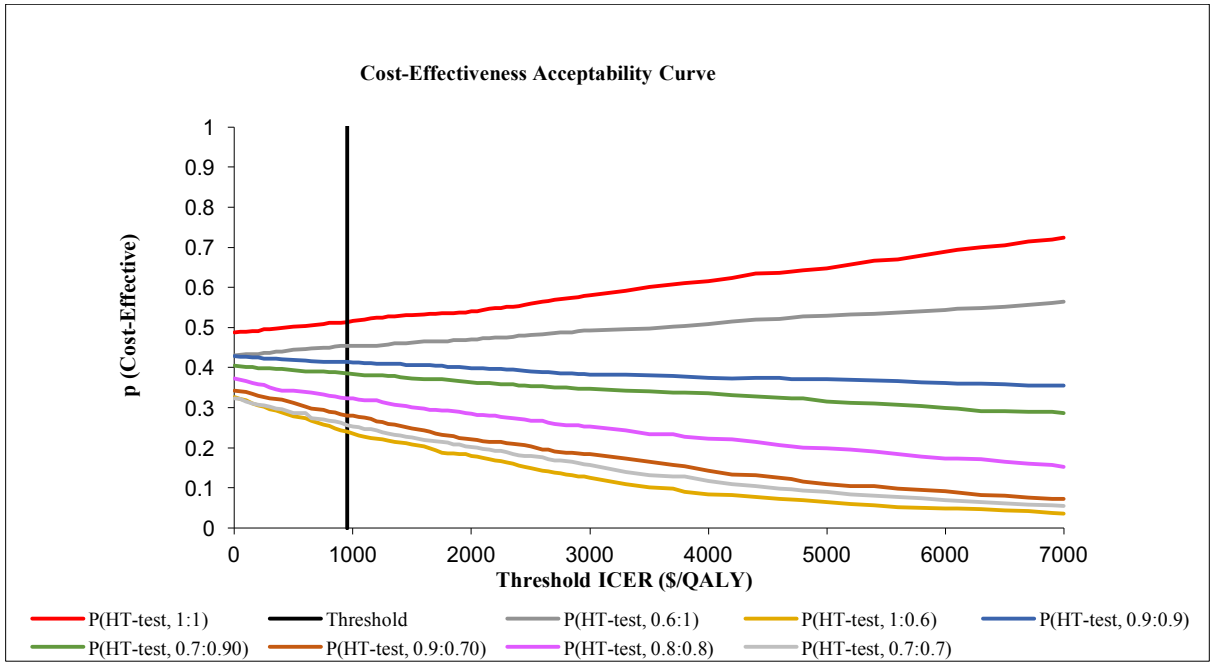
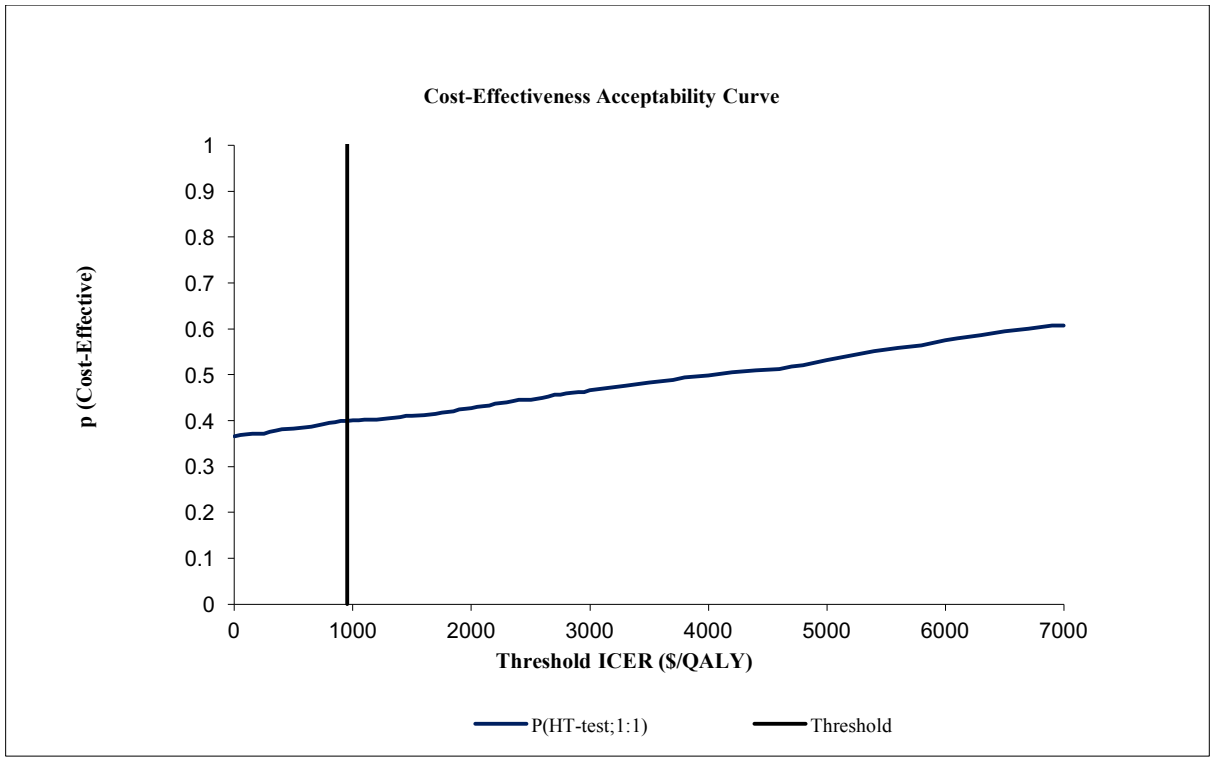


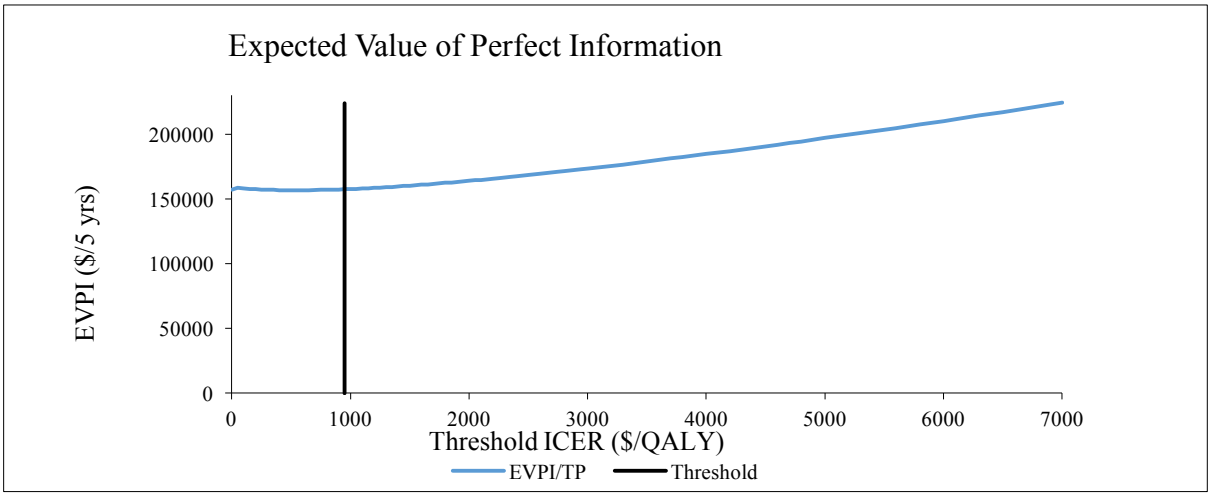
Fig A12-10 CEAC (with variations in the sensitivity and specificity of the HT-test at a fixed price of $11.00) at a WTP of $951 (Add-on role: the HT-test vs. blood culture)



Curves for the other sensitivity and specificity pairs are not shown because at a price of $11.00, only the (1,1) pair has any chance of cost-effectiveness, and even then only above the current threshold.

Appendix 13: VOI analysis (the HT-test vs. blood culture)

Fig A13-1 VOI analysis of the HT-test vs. blood culture in the replacement role at a price of $7.00



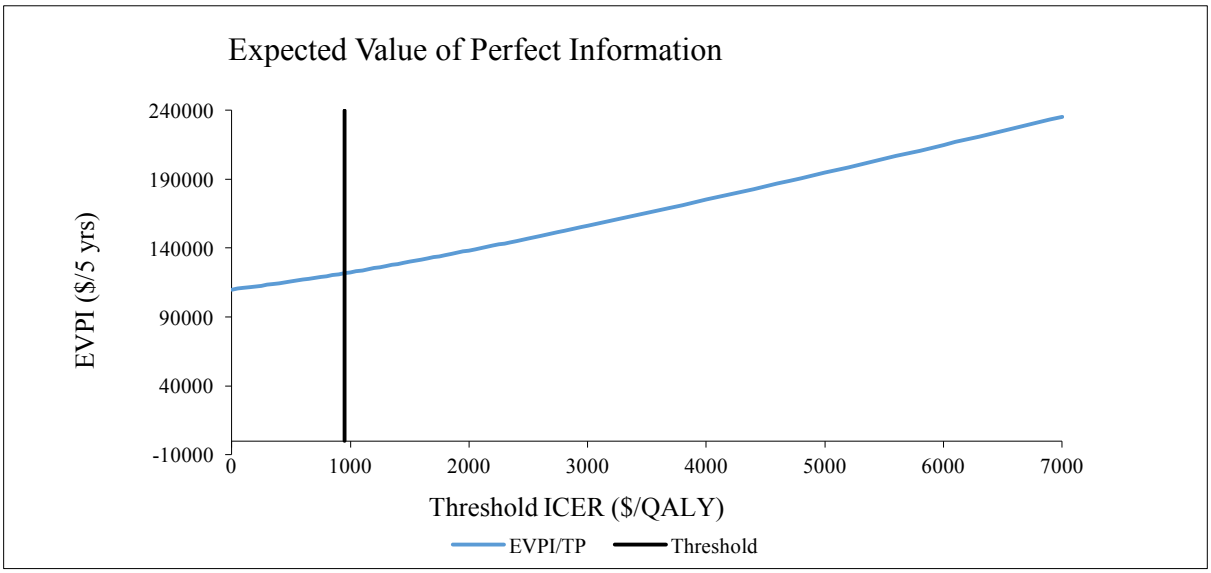
As shown in Fig A13-1, at a WTP threshold of $951/QALY, the overall EVPI for the specified decision relevance time horizon of 5 years is $157,531.

Table A13-1 Single parameter EVPPI when the HT-test is $7.00

Model parameters with uncertainty	Per Person EVPPI ($)	Standard Error	EVPPI for Ghana over 5 years ($)
Mean recovery time for successful typhoid treatment	0.00	0.00	0.0
Mean recovery time for successful malaria treatment	0.00	0.00	0.0
Prevalence of typhoid in the patient population presenting	0.150315	0.03	42146
Probability of first successful typhoid treatment	0.072177	0.02	20237
Utility when experiencing typhoid symptoms	0.00	0.00	0.0
Specificity of blood culture	0.470341	0.02	131876
Sensitivity of blood culture	0.326644	0.03	91586
Specificity of the HT-test	0.003495	0.01	980
Sensitivity of the HT-test	0.000453	0.00	127

Table A13-1 shows that the prevalence of typhoid in the patient population presenting, probability of first successful typhoid treatment, specificity and sensitivity of blood culture and the specificity and sensitivity of the HT-test are the parameters contributing to the decision uncertainty.

Fig A13-2 VOI analysis of the HT-test vs. blood culture in the triage role at a price of $6.00



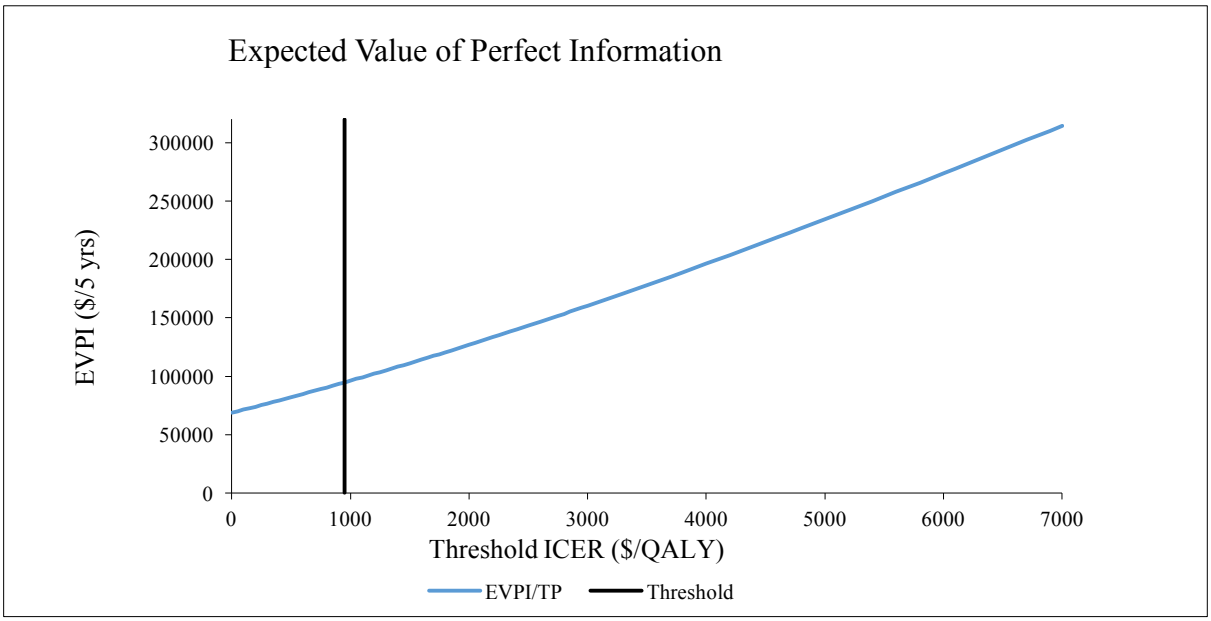
As shown in Fig A13-2, at a WTP threshold of $951/QALY, the overall EVPI for the specified decision relevance time horizon of 5 years is $121,748.

Table A13-2 Single parameter EVPPI when the HT-test is $6.00

Model parameters with uncertainty	Per Person EVPPI ($)	Standard Error	EVPPI for Ghana over 5 years ($)
Mean recovery time for successful typhoid treatment	0.00	0.00	0
Mean recovery time for successful malaria treatment	0.00	0.00	0
Prevalence of typhoid in the patient population presenting	0.236156	0.02	66214
Probability of first successful typhoid treatment	0.006789	0.01	1903
Utility when experiencing typhoid symptoms	0.00	0.00	0
Specificity of blood culture	0.263581	0.02	73904
Sensitivity of blood culture	0.186888	0.02	52400
Specificity of the HT-test	0.00	0.00	0
Sensitivity of the HT-test	0.00	0.00	0

Table A13-2 shows that the prevalence of typhoid in the patient population presenting, probability of first successful typhoid treatment and specificity and sensitivity of blood culture are the parameters contributing to the decision uncertainty.

Fig A13-3 VOI analysis of the HT-test vs. blood culture in the triage role at a price of $5.00



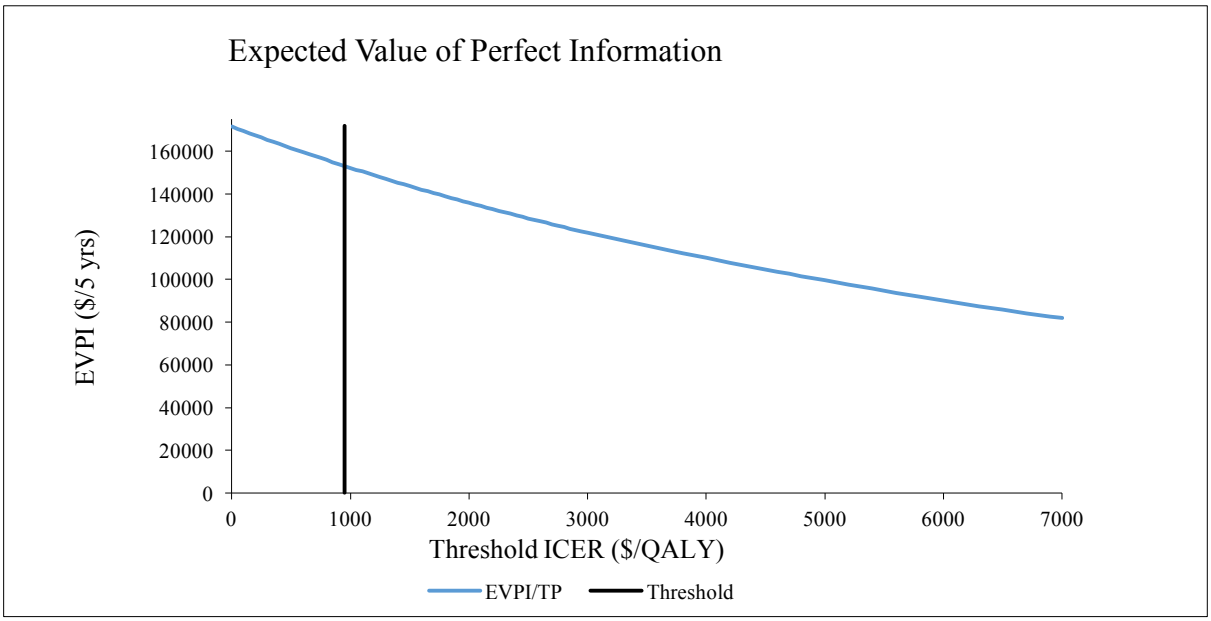
As shown in Fig A13-3, at a WTP threshold of $951/QALY, the overall EVPI for the specified decision relevance time horizon of 5 years is $94,691.

Table A13-3 Single parameter EVPPI when the HT-test is $5.00

Model parameters with uncertainty	Per Person EVPPI ($)	Standard Error	EVPPI for Ghana over 5 years ($)
Mean recovery time for successful typhoid treatment	0.00	0.00	0
Mean recovery time for successful malaria treatment	0.00	0.00	0
Prevalence of typhoid in the patient population presenting	0.161138	0.02	45180
Probability of first successful typhoid treatment	0.027984	0.01	7846
Utility when experiencing typhoid symptoms	0.00	0.00	0
Specificity of blood culture	0.142465	0.02	39945
Sensitivity of blood culture	0.118373	0.02	33190
Specificity of the HT-test	0.00	0.00	0
Sensitivity of the HT-test	0.00	0.00	0

Table A13-3 shows that the prevalence of typhoid in the patient population presenting, probability of first successful typhoid treatment and specificity and sensitivity of blood culture are the parameters contributing to the decision uncertainty.

Fig A13-4 VOI analysis of the HT-test vs. blood culture in the add-on role at a price of $10.00



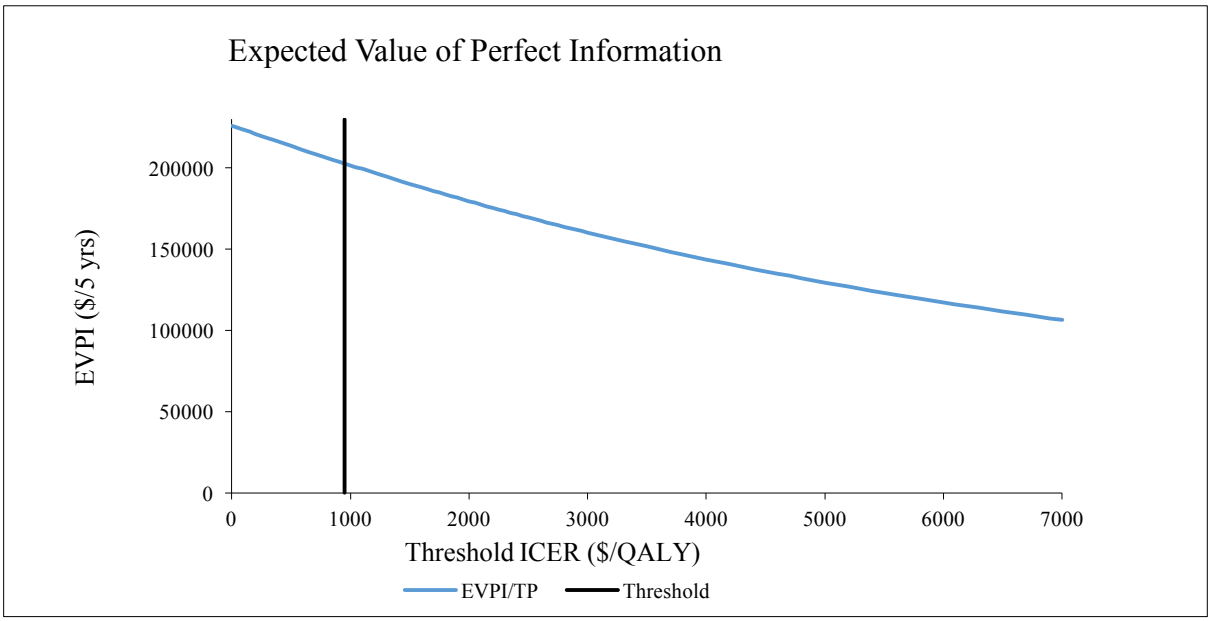
As shown in Fig A13-4, at a WTP threshold of $951/QALY, the overall EVPI for the specified decision relevance time horizon of 5 years is $153,154.

Table A13-4 Single parameter EVPPI when the HT-test is $10.00

Model parameters with uncertainty	Per Person EVPPI ($)	Standard Error	EVPPI for Ghana over 5 years ($)
Mean recovery time for successful typhoid treatment	0.00	0.00	0
Mean recovery time for successful malaria treatment	0.00	0.00	0
Prevalence of typhoid in the patient population presenting	0.234678	0.02	65800
Probability of first successful typhoid treatment	0.00	0.00	0
Utility when experiencing typhoid symptoms	0.00	0.00	0
Specificity of blood culture	0.331406	0.03	92921
Sensitivity of blood culture	0.312028	0.02	87488
Specificity of the HT-test	0.00	0.00	0
Sensitivity of the HT-test	0.00	0.00	0

Table A13-4 shows that the prevalence of typhoid in the patient population presenting and specificity and sensitivity of blood culture are the parameters contributing to the decision uncertainty.

Fig A13-5 VOI analysis of the HT-test vs. blood culture in the add-on role at a price of $9.00 (sensitivity and specificity; 0.69:0.83)



As shown in Fig A13-5, at a WTP threshold of $951/QALY, the overall EVPI for the specified decision relevance time horizon of 5 years is $202,829.

Table A13-5 Single parameter EVPPI when the HT-test is $9.00

Model parameters with uncertainty	Per Person EVPPI ($)	Standard Error	EVPPI for Ghana over 5 years ($)
Mean recovery time for successful typhoid treatment	0.00	0.01	0
Mean recovery time for successful malaria treatment	0.00	0.01	0
Prevalence of typhoid in the patient population presenting	0.412698	0.03	115714
Probability of first successful typhoid treatment	0.044505	0.03	12479
Utility when experiencing typhoid symptoms	0.00	0.00	0
Specificity of blood culture	0.494482	0.03	138645
Sensitivity of blood culture	0.476746	0.03	133672
Specificity of the HT-test	0.00	0.01	0
Sensitivity of the HT-test	0.00	0.01	0

Table A13-5 shows that the prevalence of typhoid in the patient population presenting, probability of first successful typhoid treatment and specificity and sensitivity of blood culture are the parameters contributing to the decision uncertainty.

List of references

AHRQ, (2012): Methods guide for medical test reviews. [ONLINE]. Available from http://www.google.co.uk/url?sa=t&rct=j&q=&esrc=s&frm=1&source=web&cd=4&ved=0ahUKEwjRhpbq2-rKAhUDNxQKHaPZDeoQFgg2MAM&url=http%3A%2F%2Fwww.effectivehealthcare.ahrq.gov%2Fehc%2Fproducts%2F246%2F558%2FMethods-Guide-for-Medical-Test-Reviews_Full-Guide_20120530.pdf&usg=AFQjCNHYLWVSnSvuwiHBRrX1yyNMbqIBqg&bvm=bv.113943665,d.ZWU [Accessed on 23/09/2015]

Akobeng, A.K. (2007): Understanding diagnostic tests 1: sensitivity, specificity and predictive values. **Acta paediatrica,** 96:338-341.

Akullian, A., Ng'eno, E., Matheson, A.I., Cosmas, L., Macharia, D., Fields, B., Bigogo, G., Mugoh, M., John-Stewart, G., Walson, J.L. and Wakefield, J., (2015): Environmental transmission of typhoid fever in an urban slum. **PLoS neglected tropical diseases.** 9: p.e0004212.

Alam, M.N., Haq, S.A., Das, K.K., Baral, P.K., Mazid, M.N., Siddique, R.U., Rahman, K.M., Hasan, Z., Khan, M.A.S. and Dutta, P., (1995): Efficacy of ciprofloxacin in enteric fever: comparison of treatment duration in sensitive and multidrug-resistant Salmonella. **The American journal of tropical medicine and hygiene,** 53:306-311.

Ansah, E.K., Epokor, M., Whitty, C.J., Yeung, S. and Hansen, K.S. (2013): Cost-effectiveness analysis of introducing RDTs for malaria diagnosis as compared to microscopy and presumptive diagnosis in central and peripheral public health facilities in Ghana. **The American journal of tropical medicine and hygiene,** 89**:**724-736.

Ara, R. and Brazier, J.E., (2011): Using health state utility values from the general population to approximate baselines in decision analytic models when condition-specific data are not available. **Value in Health,** 14:539-545.

Asch, D.A., Patton, J.P. and Hershey, J.C. (1990): Knowing for the Sake of Knowing: The Value of Prognostic Information**. Medical Decision Making,** 10:47-57.

Bartlett J.W and Frost, C. (2008): Reliability, repeatability and reproducibility: analysis of measurement errors in continuous variables. **Ultrasound in Obstetrics & Gynecology**, 31: 466-475.

Bertram, M.Y., Lauer, J.A., De Joncheere, K., Edejer, T., Hutubessy, R., Kieny, M.P. and Hill, S.R., (2016): Cost–effectiveness thresholds: pros and cons. **Bulletin of the World Health Organization,** 94: 925-930.

Bosch JL, Beinfeld MT, Muller JE, Brady T, Gazelle GS., (2005): A Cost–Effectiveness Analysis of a Hypothetical Catheter-Based Strategy for the Detection and Treatment of Vulnerable Coronary Plaques with Drug-Eluting Stents. **Journal of interventional cardiology**.; 18(5):339-349.

Bossuyt, P.M. and McCaffery, K. (2009): Additional patient outcomes and pathways in evaluations of testing. **Medical Decision Making**, 29: E30-E38.

Bossuyt, P.M., Irwig, L., Craig, J. and Glasziou, P. (2006): Diagnosis: Comparative Accuracy: Assessing New Tests against Existing Diagnostic Pathways. **British Medical Journal**, 332: 1089-1092.

Braun, V. and Clarke, V. (2006): Using thematic analysis in psychology. **Qualitative research in psychology**, 3:77-101.

Brennan A, Chick SE, Davies R. (2006): A taxonomy of model structures for economic evaluation of health technologies. **Health economics**. 15:1295-1310.

Briggs, A., Sculpher, M. and Claxton, K. (2006): **Decision modelling for health economic evaluation**, Oxford University Press, New York.

Buisman LR, Rutten-van Mölken MP, Postmus D, Luime JJ, Uyl-de Groot CA, Redekop WK., (2016): The early bird catches the worm: early cost-effectiveness analysis of new medical tests. **International journal of technology assessment in health care**. 3:46-53.

Buisman LR, Luime JJ, Oppe M, Hazes JM, Rutten-van Mölken MP., (2016): A five-year model to assess the early cost-effectiveness of new diagnostic tests in the early diagnosis of rheumatoid arthritis. **Arthritis research & therapy**.; 118:135.

Canh DG, Whittington D, Thoa LT, Utomo N, Hoa NT, Poulos C, Thuy DT, Kim D, Nyamete A, Acosta C., (2006): Household demand for typhoid fever vaccines in Hue, Vietnam. **Health policy and planning**. 21:241-55.

Chandey, M. and Multani, A.S., (2012): A comparative study of efficacy and safety of azithromycin and ofloxacin in uncomplicated typhoid Fever: a randomised, open labelled study. **Journal of clinical and diagnostic research.** 6:1736-1739.

Chen MH and Willan AR., (2014): Value of information methods for assessing a new diagnostic test. **Statistics in medicine**. 33:1801-1815.

Chilcott, J., Brennan, A., Booth, A., Karnon, J. And Tappenden, P. (2003): The role of modelling in prioritising and planning clinical Trials. **Health Technology Assessment**. 7:1-125.

CIA, (2015): THE WORLD FACTBOOK [ONLINE]. Available from https://www.cia.gov/library/publications/the-world-factbook/geos/gh.html [Accessed on 15/02/2016].

Cook J, Jeuland M, Whittington D, Poulos C, Clemens J, Sur D, Anh DD, Agtini M, Bhutta Z., (2008): DOMI Typhoid Economics Study Group. The cost-effectiveness of typhoid Vi vaccination programs: calculations for four urban sites in four Asian countries. **Vaccine.** 26:6305-6316.

Cook J, Sur D, Clemens J, Whittington D (2009): Evaluating investments in typhoid vaccines in two slums in Kolkata, India. **Journal of health, population, and nutrition**. 27:711.

Davenport, C.F., Takwoingi, Y., Leeflang, M.M. and Deeks, J.J. (2014): Introduction to test accuracy. Lesson 1.3: Cochrane Collaboration DTA Online Learning Materials [ONLINE]. Available from http://methods.cochrane.org/sdt/dta-author-training-online-learning [Accessed on 13/03/2016]

Deeks, J.J., Takwoingi, Y., Leeflang, M.M and Davenport C.F. (2014): Uses of Medical Tests. Lesson 1.1: Cochrane Collaboration DTA Online Learning Materials [ONLINE]. Available from http://methods.cochrane.org/sdt/dta-author-training-online-learning [Accessed on 13/03/2016]

Drain, P.K., Hyle, E.P., Noubary. F., Freedberg, K.A., Wilson, D., Bishai, W.R., Rodriguez, W. and Bassett, I.V. (2013): Diagnostic point-of-care tests in resource-limited settings. Available from: http://download.thelancet.com/pdfs/journals/laninf/PIIS1473309913702500.pdf [Accessed on 15/10/2016].

Drake TL, Devine A, Yeung S, Day NP, White LJ, Lubell Y., (2016): Dynamic transmission economic evaluation of infectious disease interventions in low- and middle- income countries: a systematic literature review. **Health economics**.25:124-139.

Drummond MF and Jefferson TO., (1996): Guidelines for authors and peer reviewers of economic submissions to the BMJ. **British Medical Journal**. 313:275-284.

Drummond, M.F., Sculpher, M.J., Torrance, G.W., O'brien, B.J. and Stoddart, G.L. (2005): **Methods for the Economic Evaluation of Health Care Programmes,** Third Edition, Oxford University Press, New York.

EUnetHTA, 2015: Methods for health economic evaluations- A guideline based on current practices in Europe https://www.eunethta.eu/wp-content/uploads/2018/03/Methods_for_health_economic_evaluations.pdf Assessed on 18/09/18

Felgner J, Jain A, Nakajima R, et al., (2017): Development of ELISAs for diagnosis of acute typhoid in Nigerian children. **PLOS Neglected Tropical Diseases**.11(6): e0005679.

Fenwick E, Palmer S, Claxton K, Sculpher M, Abrams K, Sutton A. (2006): An iterative Bayesian approach to health technology assessment: application to a policy of preoperative optimization for patients undergoing major elective surgery. **Medical Decision Making**. 26:480-496.

Ferrante di Ruffano LF, Hyde CJ, McCaffery KJ, Bossuyt PM, Deeks JJ. (2012): Assessing the value of diagnostic tests: a framework for designing and evaluating trials. **British Medical Journal.** 344: e686.

Ferrante di Ruffano LF, Dinnes J, Taylor-Phillips S, Davenport C, Hyde C, Deeks JJ. (2017): Research waste in diagnostic trials: a methods review evaluating the reporting of test-treatment interventions. **BMC medical research methodology**. 17:32.

Gale, N.K., Heath, G., Cameron, E., Rashid, S. and Redwood, S. (2013): Using the framework method for the analysis of qualitative data in multi-disciplinary health research. **BMC medical research methodology**, 13:117

Gazelle, G.S., McMahon, P.M., Siebert, U. and Beinfeld, M.T. (2005): Cost-effectiveness Analysis in the Assessment of Diagnostic Imaging Technologies 1. **Radiology**, 235:361-370.

GNDP, (2014): Standard treatment Guidelines [online]. Available from: http://collections.infocollections.org/whocountry/en/d/Js6861e/15.4.html [Accessed on 30/03/2016].

Gonzalez-Perez JG. (2002): Developing a scoring system to quality assess economic evaluations. **The European Journal of Health Economics**. 3:131-136.

Grabowski, H. (1997): The effect of Pharmacoeconomics on company research and development decisions. **Pharmacoeconomics**, 11: 389-397.

Grynberg, S., Lachish, T., Kopel, E., Meltzer, E. and Schwartz, E., (2015): Artemether-lumefantrine compared to atovaquone-proguanil as a treatment for uncomplicated Plasmodium falciparum malaria in travelers. **The American journal of tropical medicine and hygiene.** 92:13-17.

Hansen, K.S., Grieve, E., Mikhail, A., Mayan, I., Mohammed, N., Anwar, M., Baktash, S.H., Drake, T.L., Whitty, C.J., Rowland, M.W. and Leslie, T.J. (2015): Cost-effectiveness of malaria diagnosis using rapid diagnostic tests compared to microscopy or clinical symptoms alone in Afghanistan. **Malaria journal**, 14: 1-15.

Hansen KS, Ndyomugyenyi R, Magnussen P, Lal S, Clarke SE. (2017): Cost-effectiveness analysis of malaria rapid diagnostic tests for appropriate treatment of malaria at the community level in Uganda. **Health policy and planning**. 32:676-689.

Harbord, R.M. and Whiting, P., 2009. Metandi: meta-analysis of diagnostic accuracy using hierarchical logistic regression. **Stata Journal**. 9:211-229.

Hartz S, and John J. (2008): Contribution of economic evaluation to decision making in early phases of product development: a methodological and empirical review. **International Journal of Technology Assessment in Health Care**. 24:465-472.

Hartz S, and John J. (2009): Public health policy decisions on medical innovations: What role can early economic evaluation play? **Health Policy**. 89:184-192.

Huang W, Gaydos CA, Barnes MR, Jett-Goheen M, Blake DR. (2013): Comparative effectiveness of a rapid point-of-care test for detection of Chlamydia trachomatis among women in a clinical setting. **Sexually transmitted infections**. 89:108-114.

Husbands, S., Jowett, S., Barton, P. and Coast, J. (2017): How Qualitative Methods Can be Used to Inform Model Development. **PharmacoEconomics**, 35:607-612.

Husbands, S. (2016): A qualitative analysis of the model building process [ONLINE]. Available from http://etheses.bham.ac.uk/6829/ [Accessed on 01/09/2017].

Iglesias, C.P., Thompson, A., Rogowski, W.H. and Payne, K. (2016): Reporting guidelines for the use of expert judgement in model-based economic evaluations. **PharmacoEconomics**, 34: 1161-1172.

Irwig, L., Bossuyt, P., Glasziou, P., Gatsonis, C. and Lijmer, J. (2002): Designing studies to ensure that estimates of test accuracy are transferable. **British Medical Journal,** 324:669-671.

Jackson, D. (2012): HEALTHCARE ECONOMICS MADE EASY, First Edition, Scion Publishing Ltd, Oxfordshire.

Jansen, H. (2010): The logic of qualitative survey research and its position in the field of social research methods. **Forum: Qualitative Social Research** (Vol. 11, No. 2).

Johnston, K., Buxton, M.J., Jones, D.R. and Fitzpatrick, R. (1999): Assessing the costs of healthcare technologies in clinical trials [ONLINE]. Available from http://www.journalslibrary.nihr.ac.uk/__data/assets/pdf_file/0004/64831/FullReport-hta3060.pdf [Accessed on 06/02/2019]

Kaltenthaler, E.C., Dickson, R., Boland, A., Carroll, C., Fitzgerald, P., Papaioannou, D. and Akehurst, R. (2012): A qualitative study of manufacturers9 submissions to the UK NICE single technology appraisal process. **British Medical Journal open**, 2: e000562.

Kaltenthaler, E., Essat, M., Tappenden, P. and Paisley, S. (2014): Identification and review of cost-effectiveness model parameters: a qualitative study. **International journal of technology assessment in health care**, 30:333-340.

Keddy, K.H., Sooka, A., Letsoalo, M.E., Hoyland, G., Chaignat, C.L., Morrissey, A.B. and Crump, J.A. (2011): Sensitivity and specificity of typhoid fever rapid antibody tests for laboratory diagnosis at two sub-Saharan African sites. **Bulletin of the World Health Organization**, 89:640-647.

Keeling MJ, and Rohani P. (2011): Modeling infectious diseases in humans and animals. Princeton University Press.

Kelley, K., Clark, B., Brown, V. and Sitzia, J. (2003): Good practice in the conduct and reporting of survey research. **International Journal for Quality in health care**, 15:261-266.
Kind, P., Lafata, J.E., Matuszewski, K. and Raisch, D. (2009): The Use of QALYs in Clinical and Patient Decision-Making: Issues and Prospects. **VALUE IN HEALTH,** 12: 27-30.

Kind, P., Lafata, J.E., Matuszewski, K. and Raisch, D. (2009): The Use of QALYs in Clinical and Patient Decision-Making: Issues and Prospects. **VALUE IN HEALTH,** 12: 27-30.

Kip, M.M., IJzerman, M.J., Henriksson, M., Merlin, T., Weinstein, M.C., Phelps, C.E., Kusters, R. and Koffijberg, H., (2018): Toward Alignment in the Reporting of Economic Evaluations of Diagnostic Tests and Biomarkers: The AGREEDT Checklist. **Medical decision making**, 38:778-788.

Knottnerus, J.A., Van Weel, C. and Muris, J.W. (2002): Evaluation of diagnostic procedures. **British Medical Journal,** 324:477-480.

Koffijberg, H., van Zaane, B. and Moons, K.G. (2013): From accuracy to patient outcome and cost-effectiveness evaluations of diagnostic tests and biomarkers: an exemplary modelling study. **BMC medical research methodology**, 13:1-12.

Lauria DT, Maskery B, Poulos C, Whittington D. (2009): An optimization model for reducing typhoid cases in developing countries without increasing public spending. **Vaccine**. 27:1609-1621.

Lee, D.W., Neumann, P.J. and Rizzo, J.A. (2010): Understanding the medical and nonmedical value of diagnostic testing. **Value in health**, 13:310-314.

Leeflang, M.M., Deeks, J.J., Gatsonis, C. and Bossuyt, P.M. (2008): Systematic reviews of diagnostic test accuracy. **Annals of internal medicine**, 149:889-897.

Leeflang, M.M., Deeks, J.J., Takwoingi, Y. and Macaskill, P. (2013): Cochrane diagnostic test accuracy reviews. **Systematic reviews**, 2: 1-6.

Lin, P.J., Cangelosi, M.J., Lee, D.W. and Neumann, P.J. (2013): Willingness to pay for diagnostic technologies: a review of the contingent valuation literature. **Value in Health**. 16:797-805.

Lo, N.C., Gupta, R., Stanaway, J.D., Garrett, D.O., Bogoch, I.I., Luby, S.P. and Andrews, J.R., (2018): Comparison of Strategies and Incidence Thresholds for Vi Conjugate Vaccines Against Typhoid Fever: A Cost-effectiveness Modeling Study. **The Journal of infectious diseases**. S1-S11.

Lord, S.J., Irwig, L. and Bossuyt, P.M. (2009): Using the principles of randomized controlled trial design to guide test evaluation. **Medical Decision Making**, 29: E1-E12.

Lugnér, A.K., Mylius, S.D. and Wallinga, J., (2010): Dynamic versus static models in cost-effectiveness analyses of anti-viral drug therapy to mitigate an influenza pandemic. **Health economics**, 19:518-531.

McCabe, C., Claxton, K. and Culyer, A.J., 2008. The NICE cost-effectiveness threshold. What it is and what it means. **Pharmacoeconomics**, 26:733-744.

Mallma, P., Garcia, P., Carcamo, C., Torres-Rueda, S., Peeling, R., Mabey, D. and Terris-Prestholt, F. (2016): Rapid Syphilis Testing Is Cost-Effective Even in Low-Prevalence Settings: The CISNE-PERU Experience**. PloS one**, 11: e0149568.

Markiewicz K, van Til JA, IJzerman MJ. (2014): Medical devices early assessment methods: systematic literature review. **International journal of technology assessment in health care**. 30:137-146.

Marks, F., Adu-Sarkodie, Y., Hünger, F., Sarpong, N., Ekuban, S., Agyekum, A., Nkrumah, B., Schwarz, N.G., Favorov, M.O., Meyer, C.G. and May, J. (2010): Typhoid fever among children, Ghana. **Emerging Infectious Diseases**, 16: 176-178.

Matchar, D.B. (2012): Introduction to the methods guide for medical test reviews. **Journal of general internal medicine,** 27:4-10.

McAteer H, Cosh E, Freeman G, Pandit A, Wood P, Lilford R. (2007): Cost-effectiveness analysis at the development phase of a potential health technology: examples based on tissue engineering of bladder and urethra. **Journal of tissue engineering and regenerative medicine**. 1:343-349.

Merlin, T., Lehman, S., Hiller, J.E. and Ryan, P. (2013): The "linked evidence approach" to assess medical tests: a critical analysis. **International journal of technology assessment in health care,** 29: 343-350.

Merlin, T., 2014. The use of the 'linked evidence approach'to guide policy on the reimbursement of personalized medicines. **Personalized Medicine**, 11: 435-448.

Morris, S., Devlin, N., Parkin, D. and Spencer, A. (2012): **Economic Analysis in Health Care,** Second Edition, John Wiley and Sons, West Sussex.

MSAC, (2005): Guidelines for the assessment of diagnostic technologies. [ONLINE]. Available fromhttp://www.google.co.uk/url?sa=t&rct=j&q=&esrc=s&frm=1&source=web&cd=4&ved=0ahUKEwiMnP3rzerKAhXE0xQKHXneBugQFggzMAM&url=http%3A%2F%2Fwww.health.gov.au%2Finternet%2Fmsac%2Fpublishing.nsf%2FContent%2FD81BE529B98B3DB6CA2575AD0082FD1B%2F%24File%2FGuidelines%2520for%2520the%2520assessment%2520of%2520diagnostic%2520technologies%2520Sept%25202005.pdf&usg=AFQjCNH0foiLaBNtn690xCdzMv3SUQrHDw&bvm=bv.113943665,d.d24.[Accessed on 23/11/2015]

Murray, C.J., Vos, T., Lozano, R, et al (2012): Disability-adjusted life years (DALYs) for 291 diseases and injuries in 21 regions, 1990-2010: a systematic analysis for the Global Burden of Disease Study 2010. **The lancet**, 380:2197-2223.

Musgrove P. (1992): Cost-benefit analysis of a regional system for vaccination against pneumonia, meningitis type B, and typhoid fever. **Bull PAHO**. 26: 173-191.

Mutua, J.M., Wang, F.B. and Vaidya, N.K., 2015. Modeling malaria and typhoid fever co-infection dynamics. **Mathematical biosciences**. 264:128-144.

Neil, K.P., Sodha, S.V., Lukwago, L., Shikanga, O., Mikoleit, M., Simington, S.D., Mukobi, P., Balinandi, S., Majalija, S., Ayers, J. and Kagirita, A. (2012): A large outbreak of typhoid fever associated with a high rate of intestinal perforation in Kasese District, Uganda, 2008–2009. **Clinical Infectious Diseases**, 54: 1091-1099.

Neumann, P.J., Cohen, J.T., Hammitt, J.K., Concannon, T.W., Auerbach, H.R., Fang, C. and Kent, D.M. (2012): Willingness-to-pay for predictive tests with no immediate treatment implications: a survey of US residents. **Health economics**, 21:238-251.

NICE, (2011): **Diagnostics Assessment Programme manual.** [ONLINE]. Available from: http://www.google.co.uk/url?sa=t&rct=j&q=&esrc=s&frm=1&source=web&cd=1&ved=0ahUKEwirjvTuyurKAhVJQBQKHe77CS8QFggcMAA&url=http%3A%2F%2Fwww.nice.org.uk%2FMedia%2FDefault%2FAbout%2Fwhat-we-do%2FNICE-guidance%2FNICE-diagnostics-guidance%2FDiagnostics-assessment-programme-manual.pdf&usg=AFQjCNEi6_ApSXVuG2cah8Oz1oYXNVh5nA&bvm=bv.113943665,d.bGs [Accessed on 23/11/2015].

Nonvignon, J., Aryeetey, G.C., Malm, K.L., Agyemang, S.A., Aubyn, V.N., Peprah, N.Y., Bart-Plange, C.N. and Aikins, M., (2016): Economic burden of malaria on businesses in Ghana: a case for private sector investment in malaria control. **Malaria journal**, 15: p.454.

Nosten, F. and White, N.J., (2007): Artemisinin-based combination treatment of falciparum malaria. **The American journal of tropical medicine and hygiene**. 77:181-192.

Odame, E.A. (2013): Systematic Review of Economic Evaluation Literature in Ghana: Is Health Technology Assessment the Future? **Value in Health Regional Issues**, 2:279-283.

Parikh, R., Mathai, A., Parikh, S., Sekhar, G.C. and Thomas, R. (2008): Understanding and using sensitivity, specificity and predictive values. **Indian journal of ophthalmology**. 56:1-17.

Pauker, S.G. and Kassirer, J.P. (1987): Decision analysis. **New England Journal of Medicine.** 316: 250-258.

Peeling, R.W., Smith, P.G. and Bossuyt, P.M., (2008): A guide for diagnostic evaluations**.** **Nature Reviews Microbiology**, 8: S2-S6.

Petrou, S and Gray, A. (2011): Economic evaluation alongside randomised controlled trials: design, conduct, analysis and reporting. **British Medical Journal**, 342: d1548.

Philips Z, Bojke L, Sculpher M, Claxton K, Golder S. (2006): Good practice guidelines for decision-analytic modelling in health technology assessment. **Pharmacoeconomics**. 24:355-371.

Pietzsch, J.B. and Paté-Cornell, M.E. (2008): Early technology assessment of new medical devices. **International Journal of Technology Assessment in Health Care**, 24:36-44.

Pope, C and Mays, N. (2006): **Qualitative research in healthcare,** Third Edition, Blackwell publishing, Massachusetts.

Pope, C., Ziedland, S. and Mays, N., 2000. Qualitative research in health care: Analysing qualitative data. 320. **British Medical Journal.** 320:114.

Poulos C, Bahl R, Whittington D, Bhan MK, Clemens JD, Acosta CJ. (2004): Cost-benefit analysis of typhoid fever immunization programmes in an Indian Urban Slum Community. **Journal of Health, Population and Nutrition**. 22:311-321.

RAND Corporation (2007): Estimating the global health impact of improved diagnostic tools for the developing world. Available from http://www.rand.org/content/dam/rand/pubs/research_briefs/2007/RAND_RB9293.pdf [Accessed Jan 14, 2016].

Redekop, K. and Mikudina, B. (2013): Early medical technology assessments of medical devices and tests. **Journal of Health Policy and Outcomes Research**, 1: 23-37.

Redekop, W.K. (2006): Cost–effectiveness analyses of diagnostic strategies: a literature survey using the NHS Economic Evaluation Database. **Expert review of pharmacoeconomics & outcomes research**, 6:41-48.

Ritchie, J and Lewis, J. (2003): **Qualitative research practice: A guide for social science students and researchers**. Sage: London.

Roberts, M., Russell, L.B., Paltiel, A.D., Chambers, M., McEwan, P. and Krahn, M. (2012): Conceptualizing a model: A report of the ISPOR-SMDM modeling good research practices task force–2. **Medical Decision Making**, 32: 678-689.

Schito, M.L., Peter, T.F., Cavanaugh, S., Piatek, A.S., Young, G.J., Alexander, H., Coggin, W., Domingo, G.J., Ellenberger, D., Ermantraut, E. and Jani, I.V. (2012): Opportunities and challenges for cost-efficient implementation of new point-of-care diagnostics for HIV and tuberculosis. **Journal of Infectious Diseases**, 205:1-12.

Schroeder, L.F., Elbireer, A., Jackson, J.B. and Amukele, T.K. (2015): Laboratory Diagnostics Market in East Africa: A Survey of Test Types, Test Availability, and Test Prices in Kampala, Uganda. **PloS one,** 10:1-14.

Schwartz, L.M., Woloshin, S., Fowler Jr, F.J. and Welch, H.G. (2004): Enthusiasm for cancer screening in the United States. **The Journal of the American Medical Association**, 291:71-78.

Sculpher M, Drummond M, Buxton M. (1997): The iterative use of economic evaluation as part of the process of health technology assessment. **Journal of Health Services Research**. 2:26-30.

Shepard DS, Walsh JA, Kleinau E, Stansfield S, Bhalotra S. (1995): Setting priorities for the Children's Vaccine Initiative: a cost-effectiveness approach. **Vaccine**. 13:707-14.

Sodzi-Tettey, S., Aikins, M., Awoonor-Williams, J.K. and Agyepong, I.A., (2012): Challenges in provider payment under the Ghana National Health Insurance Scheme: a case study of claims management in two districts. **Ghana medical journal**, 46:189-199.

Squires, H., Chilcott, J., Akehurst, R., Burr, J. and Kelly, M.P. (2016): A framework for developing the structure of public health economic models. **Value in Health**, 19:588-601.

Srivastava, A. and Thomson, S.B. (2009): Framework analysis: a qualitative methodology for applied policy research. **Journal of administration and governance**, 4: 72-79.

Stockham, S.L. and Scott, M.A. (2013): **Fundamentals of veterinary clinical pathology**, Second Edition, John Wiley & Sons.

Storey, H.L., Huang, Y., Crudder, C., Golden, A., de los Santos, T. and Hawkins, K., (2015): A meta-analysis of typhoid diagnostic accuracy studies: a recommendation to adopt a standardized composite reference. **PloS one.** 10: p.e0142364.

Strong, M., Oakley, J.E. and Brennan, A., (2014): Estimating multiparameter partial expected value of perfect information from a probabilistic sensitivity analysis sample: a nonparametric regression approach. **Medical Decision Making**, 34:311-326.

Sullivan, W. and Payne, K. (2011): The appropriate elicitation of expert opinion in economic models. **Pharmacoeconomics**, 29: 455-459.

Tappenden P, Chilcott JB, Eggington S, Oakley J, McCabe C. (2004): Methods for expected value of information analysis in complex health economic models: developments on the health economics of interferon-β and glatiramer acetate for multiple sclerosis. **Health Technology Assessment**. 8(27).

Tawiah T, Hansen KS, Baiden F, Bruce J, Tivura M, Delimini R, Amengo-Etego S, Chandramohan D, Owusu-Agyei S, Webster J. (2016): Cost-Effectiveness Analysis of Test-Based versus Presumptive Treatment of Uncomplicated Malaria in Children under Five Years in an Area of High Transmission in Central Ghana. **PloS one**. 11(10):e0164055.

Thornbury, J.R., Fryback, D.G. and Edwards, W. (1975): Likelihood Ratios as a Measure of the Diagnostic Usefulness of Excretory Urogram Information. **Radiology**. 114:561-565.

Tilahun GT, Makinde OD, Malonza D. (2017): Modelling and Optimal Control of Typhoid Fever Disease with Cost-Effective Strategies. **Computational and mathematical methods in medicine**. 2017:1-16.

Trikalinos TA, Siebert U, Lau J.(2009): Decision-analytic modeling to evaluate benefits and harms of medical tests: uses and limitations. **Medical Decision Making**. 2009 Sep;29(5): E22-29.

Tu HA, Han D, Pechlivanoglou P, Miller F, Rac V, Chin W, Tricco AC, Paulden M, Bielecki J, Krahn M. (2014): Early economic evaluation of emerging health technologies: protocol of a systematic review. **Systematic reviews**. 3:81.

Tuffaha HW, Gordon LG, Scuffham PA. (2014): Value of information analysis in healthcare: a review of principles and applications. **Journal of medical economics**. 17(6):377-383.

Urdea, M., Penny, L.A., Olmsted, S.S, et al. (2006): Requirements for high impact diagnostics in the developing world. **Nature**, 444: 73–79.

Vaidya A, Joore MA, ten Cate-Hoek AJ, Cate HT, Severens JL. (2014): Cost–effectiveness of risk assessment and tailored treatment for peripheral arterial disease patients. **Biomarkers in medicine**. 8:989-999.

Vallejo-Torres L, Steuten LM, Buxton MJ, Girling AJ, Lilford RJ, Young T. (2008): Integrating health economics modeling in the product development cycle of medical devices: a Bayesian approach. **International journal of technology assessment in health care**. 24:459-464.

Van den Bruel, A., Cleemput, I., Aertgeerts, B., Ramaekers, D. and Buntinx, F. (2007): The evaluation of diagnostic tests: evidence on technical and diagnostic accuracy, impact on patient outcome and cost-effectiveness is needed. **Journal of clinical epidemiology**, 60:1116-1122.

van Voorn, G.A., Vemer, P., Hamerlijnck, D., Ramos, I.C., Teunissen, G.J., Al, M. and Feenstra, T.L., (2016): The Missing Stakeholder Group: why patients should be involved in health economic modelling. **Applied health economics and health policy**. 14:129-133.

Wain, J., Hendriksen, R.S., Mikoleit, M.L., Keddy, K.H. and Ochiai, R.L (2015): Typhoid fever. **Lance**t, 385: 1136-1145.

Watson CH, and Edmunds WJ. (2015): A review of typhoid fever transmission dynamic models and economic evaluations of vaccination. **Vaccine**. 33:C42-54.

Webster, J.G. (2014): **The Physiological Measurement Handbook**, First Edition, CRC Press, Florida.

Whiting, P.F., Davenport, C., Jameson, C., Burke, M., Sterne, J.A., Hyde, C. and Ben-Shlomo, Y. (2015): How well do health professionals interpret diagnostic information? A systematic review. **BMJ** open, 5: p.e008155.

WHO, (2000): Workbook 8: Economic Evaluations [ONLINE]. Available from http://www.google.co.uk/url?sa=t&rct=j&q=&esrc=s&source=web&cd=9&sqi=2&ved=0CGIQFjAI&url=http%3A%2F%2Fwww.emcdda.europa.eu%2Fattachements.cfm%2Fatt_5870_EN_8_economic_evaluations.pdf&ei=AUrhUfCUDbO10QW1yYD4CA&usg=AFQjCNGAY80QfHJQXY5ehTAQynjSMAwDAA&bvm=bv.48705608,d.d2k [Accessed on 6/07/13].

WHO, (2003): Background document: the diagnosis, treatment and prevention of typhoid fever.

WHO, (2005): Good clinical diagnostic practice: A guide for clinicians in developing countries to the clinical diagnosis of disease and to making proper use of clinical diagnostic services [ONLINE]. Available from

http://www.google.co.uk/url?sa=t&rct=j&q=&esrc=s&frm=1&source=web&cd=2&ved=0ahUKEwiE_eX-te3KAhUJvxQKHS1zCh4QFggmMAE&url=http%3A%2F%2Fwww.emro.who.int%2Fdsaf%2Fdsa236.pdf&usg=AFQjCNEWLK93c-8KUh5oOEJ5gZVRIXYu4g [Accessed on 27/01/2016].

Wilson, J.M.G. and Jungner, G. (1968): Principles and practice of screening for disease. **World Health Organization, Public Health Paper**, (34).

Woods, B., Revill, P., Sculpher, M. and Claxton, K., (2016): Country-level cost-effectiveness thresholds: initial estimates and the need for further research. **Value in Health**. 19:929-935.

Wu AC, Gay C, Rett MD, Stout N, Weiss ST, Fuhlbrigge AL. (2015): Pharmacogenomic test that predicts response to inhaled corticosteroids in adults with asthma likely to be cost-saving. **Pharmacogenomics**. 16:591-600.