Enhanced Decision Models for the Diagnosis and Treatment of Malaria in an age of ACTs

Yoel Lubell

Thesis submitted towards the degree of Doctor of Philosophy at the University of London

Health Policy Unit

London School of Hygiene and Tropical Medicine

University of London

January 2009

Statement of work

The data for the analyses in this thesis were obtained through collaborations with individuals and organizations carrying out four randomized control trials in Africa and Asia. My role in these trials was to assist in designing their economic evaluation with the advice of LSHTM staff, and to carry out the subsequent analyses. These studies form the core of the thesis. In the course of work on the thesis the following articles were also prepared and published:

1) Lubell, Y., H. Reyburn, et al. (2007). "The cost-effectiveness of parasitological diagnosis for malaria-suspected patients in an era of combination therapy." Am J Trop Med Hyg **77**(Suppl 6): 128-132.

2) Lubell, Y., H. Reyburn, et al. (2008). "The impact of response to the results of diagnostic tests for malaria: cost-benefit analysis." BMJ **336**(7637): 202-5.

3) Lubell, Y., H. Hopkins, et al. (2008). "An interactive model for the assessment of the economic costs and benefits of different rapid diagnostic tests for malaria." Malar J **7**: 21.

4) Lubell, Y., Yeung, S. et al. (In press) "Cost-effectiveness of artesunate for the treatment of severe malaria" TMIH

I certify being the lead author in the above publications, designing the models and drafting the manuscripts. Fellow authors had inputs relating to trial clinical outcomes, model structure, and reviewing and editing the manuscripts.

I have read and understood the School's definition of plagiarism and cheating given in the Research Degrees Handbook. I declare that this thesis is my own work, and that I have acknowledged all results and quotations from the published or unpublished work of other people.

Yoel Lubell   **Confirmed by Anne Mills**   **December 10th 2008**

Abstract

New diagnostics and treatments for malaria have renewed hope in the developing world as they promise relief from the debilitating effects of this illness. Accompanying these interventions are a growing number of economic evaluations assessing their efficiency. To ensure the relevance of economic evaluations to decision making purposes it is imperative that they use best available computational and statistical approaches.

This thesis initially discusses the necessary requirements for economic evaluations to ensure they provide appropriate decision recommendations. This is followed by four evaluations of malaria diagnostics and treatments using methods new to the context of malaria.

The first study expands the range of factors included in the evaluation of diagnostic tests, addressing compromised adherence to test results and societal costs associated with antimalarial use.

The second analysis demonstrates how models can be designed as decision support tools allowing stakeholders to enter local data along with other parameter estimates, priorities and values. Both Bayesian and deterministic models are presented for comparison.

The third analysis demonstrates the use of multilevel models for economic evaluations based on multi-centre trials. The chapter compares the results of a multilevel model evaluating treatments for severe malaria with those obtained in a standard analysis.

The fourth study uses a Markov model to evaluate the efficiency of Home Management of Malaria programmes. The use of a Markov model addresses the restricted portrayal of malaria infection and illness that has characterised many previous evaluations.

In addition to contributing to a better understanding of the cost-effectiveness of the latest malaria treatments and diagnostic tests, this thesis seeks to bridge the growing gap between recent methodological advances in the field of economic evaluation, and the relative paucity of evaluations producing practical and effective policy recommendations for areas of the world where the burden of malaria and other diseases is heaviest.

Table of Contents

List of Figures

List of Tables

List of abbreviations

ACT - Artemisinin Combination Therapy

AL – Artemether-Lumefantrine

CBA – Cost-benefit Analysis

CQ - Chloroquine

CEA – Cost-effectiveness Analysis

CDDs – Community Drug Distributors

CUA – Cost-utility analysis

DA – Decision Analysis

DALY – Disability Adjusted Life Year

DST – Decision Support Tool

HoT – Harm of Treatment

ICER – Incremental Cost-effectiveness Ratio

ITN – Insecticide Treated Nets

MLM – Multi-level Modelling

MoH – Ministry of Health

QALY – Quality Adjusted Life Year

RDT - Rapid Diagnostic Test

SP - Sulfadoxine-Pyrimethamine

SSA – Sub Saharan Africa

WHO – World Health Organization

WTP – Willingness to Pay

YLL – Year of Life Lost

Acknowledgements

"The first lesson of economics is scarcity: There is never enough of anything to satisfy all those who want it. The first lesson of politics is to disregard the first lesson of economics."

Thomas Sowell

Preface

This thesis is the product of three years' work evaluating several malaria diagnostics and treatments from an economic perspective. At the start of this period a number of areas in the economic evaluation of malaria treatment and diagnostics that needed improvement were suggested by health economists at LSHTM. Other issues were identified by reviewing the existing literature. To address these issues, several potentially useful methods not previously applied to malaria interventions were selected. As these related to different interventions and contexts, it was decided to pursue these in individual analyses. Data for these analyses were obtained by collaborating with investigators carrying out clinical trials of antimalarials, rapid diagnostic tests, and distribution strategies. The main body of the thesis is comprised of these studies.

While the individual analyses differ in the interventions they assess and the methods they use, they all represent enhancements to the economic evaluation of malaria interventions by reflecting two common themes. First, economic evaluations should be as comprehensive as possible, embracing factors beyond immediate costs and consequences to ensure that they are relevant to decision making purposes. Second, evaluations should ensure that their results are as responsive as possible to variation in local circumstances.

Readers well versed in economic evaluations in the context of developed countries will be familiar with many of the methods used in these analyses. In the context of malaria however, economic evaluation and decision making practices have lagged far behind those used in the developed world, despite the colossal health burden malaria places on much of the population. This thesis contributes to knowledge by drawing on methods developed recently in high income countries and applying them to decision making related to malaria.

Thesis structure

The thesis is divided into two sections. The first section, comprised of four chapters, provides an overview of the past and present dilemmas policy makers face regarding the use of malaria treatment and diagnostics; it then describes different approaches decision makers take in addressing such dilemmas. Chapter 2 reviews the literature of economic evaluations previously carried out to appraise malaria diagnostics and treatments, and their relevance to decision making purposes with the aim of identifying the areas for improvement. Chapter 3 explores the theoretical underpinnings of economic evaluation frameworks with a focus on those relevant for use in the context of malaria diagnosis and treatment. Chapter 3 goes on to explore different approaches to handling uncertainty in decision models and the incorporation of data into the analyses. Drawing on Chapter 3, Chapter 4 provides a presentation of the aims, objectives and methods of the studies that form the core of the thesis.

In the second section, four independent analyses demonstrate different decision models, all representing methods that are new in the context of malaria treatment and diagnostics. The four studies address some of the most urgent questions that policy makers face in relation to malaria control. Each of these studies deals with different issues; they are presented in the order in which they were conducted and to a certain extent build on each others' methodologies.

The concluding chapter discusses the re-orientation of economic evaluations to decision making requirements. The chapter argues for using advanced modelling and information-technology capacities in the malaria context to enhance the relevance of economic evaluations in particular locations, circumstances and preferences. The limitations of the methods used in the thesis are also acknowledged followed by the final conclusions and suggestions for further research.

1. Malaria, research and decision making

1.1 Introduction

This thesis focuses on the development of economic models to inform decision making regarding malaria diagnostic and treatment strategies. This chapter opens with a historical overview of the interactions between malaria research and policy making highlighting the current dilemmas concerning the adoption of the most recently developed classes of drugs and diagnostics. The focus is particularly, but not exclusively, on Sub Saharan Africa (SSA) where the burden of malaria is highest. The focus is also uniquely on interventions targeting *P.falciparum* malaria that is responsible for almost all malaria related mortality. This is followed by a brief summary of malaria pathology, epidemiology, and the general burden of malaria on the economy. Given the focus on practical decision models in this thesis, the subsequent section reviews different approaches to decision making to assess how policy makers can make use of such models when considering the adoption of new health care interventions.

1.2 A historical overview of malaria research and policy

The vast majority of the estimated one million annual deaths from malaria occur amongst African children (Breman, Alilio et al. 2004). Pregnant women are also highly susceptible to the parasite, as are older children and adults in epidemic prone areas. In addition to direct morbidity and mortality, malaria's adverse impact permeates societies through a variety of mechanisms, feeding into vicious cycles of poverty and poor health (Breman 2001). This has been observed at the micro-level by identifying factors such as impaired productivity and household health expenditure on malaria prevention and treatment (Chima, Goodman et al. 2003). It has been seen at the macro-level by correlating malaria endemicity and poor national economic performance (Sachs and Malaney 2002; Malaney, Spielman et al. 2004).

20

In much of SSA, malaria is almost synonymous with disease and poor health, with vast amounts of scarce resources being used to mitigate its effects. Households across the continent spend a considerable part of their income on malaria prevention and treatments (Chima, Goodman and Mills 2003). Many African governments devote a significant proportion of highly constrained health budgets to malaria control. At the international level unprecedented funding is being made available to both malaria related research and for its control of the disease, funding which could also be used for alternative health needs (Waddington, Martin et al. 2005).

Notwithstanding the numerous ebbs and flows, malaria has been at the centre of health research and policy in SSA and elsewhere for over a century. This period can be divided into three distinct phases defined by two elements: (1) the overall approach to malaria control; and (2) the interactions between research and decision making (Gilles and Lucas 1998; Alilio, Bygbjerg et al. 2004).

1.2.1 Early discoveries

In the late 19th and early 20th centuries, the fundamentals of malaria pathology, transmission and control were discovered. The presence of *Plasmodium* parasites in febrile patients was first detected by Laveran in 1880 (Gilles and Lucas 1998), though establishing that mosquitoes were the vectors for the parasites took another two decades and can be attributed to either Grassi or Ross – a matter of unresolved contention. The isolation of DDT was achieved in the 1880s although it was only recognized as an insecticide by Paul Muller in 1934. Chloroquine (CQ), the mainstay of malaria treatment for decades to follow, was first created in Germany in 1936 under the name of Sontochin, and later re-branded and mass produced in the USA in 1946 (White 1992).

With respect to research and development, these remained the most significant events until the latter decades of the 20th century. In implementing these technologies during this phase, it was often the same pioneers in research, such as Ronald Ross, whom also led the way operationally in attempts to control the disease (Alilio, Bygbjerg et al. 2004). In pre-independence Africa, control of the disease was attempted only on a limited scale and

mostly where it was in the interest of the colonial powers for military and commercial purposes (Gilles and Lucas 1998, White 1992).

1.2.2 Centrally run vertical programmes

Building on these discoveries, the middle of the 20th century saw a proliferation of vertical programmes in blanket applications of DDT and CQ for eradication of the disease that proved successful in many temperate climates where transmission was seasonal. Leading these programmes were national governments functioning both independently or in collaboration with the World Health Organization (WHO). The most concerted effort was made with the Malaria Eradication Programme between 1956 and 1969.

While control efforts proved successful in Europe, the US and parts of the Mediterranean, other areas such as South Asia, notably India and Sri Lanka, initially saw good outcomes but soon suffered large setbacks (Kager 2002). SSA was excluded altogether from the WHO eradication programme as it was deemed unprepared. Nevertheless, the WHO did pursue a limited eradication program in Zimbabwe, South Africa and Ethiopia where it was deemed feasible (Alilio, Bygbjerg et al. 2004). These few attempts failed and it was quickly recognised that it was too ambitious an objective for SSA and resources would be better used for control of malaria, primarily through rapid detection and treatment of patients suspected to be suffering from malaria.

The availability of CQ, an effective, safe and cheap antimalarial, was conducive to partial achievement of this aim. The drug was disseminated widely within and beyond the formal health sector, allowing patients with fever and other symptoms suggestive of malaria access to affordable and effective treatment. The dissemination of CQ was further accelerated with the focus on primary health care, spurred by the Alma Ata convention in 1978.

Consequently, CQ was readily used to treat all malaria suspected patients. Furthermore, the antipyretic characteristic of the drug alleviated symptoms for other non-malarial causes of illness. This additional benefit coincided well with the promotion of presumptive treatment strategies, where in the absence of an alternative obvious cause, all febrile

illnesses were treated as malaria. From the decision makers' perspective, CQ was the ideal drug with almost no apparent down-side to its use.

With the availability of DDT and CQ, two relatively simple and effective interventions, and the previous success with eradicating malaria in many parts of the world, research into malaria during this period fell into a lull. However, the golden age of these two interventions would soon end, as DDT was largely abandoned and CQ became ineffective, causing a surge in morbidity and mortality amongst what became a highly susceptible population (Trape 2001).

DDT was the first to face a decline in its effectiveness, as some of the mosquito vectors developed resistance to its toxicity. Furthermore, uproar over environmental concerns resulted in DDT being phased out long before it was ever deployed on a large scale in SSA - notwithstanding a number of promising trials demonstrating its potential role in reducing malaria transmission (Curtis and Lines 2000). The golden age for CQ came to an end with the spread of parasites resistant to its therapeutic powers, initially in Southeast Asia, and later in most parts of Africa. Eventually CQ resistance spread to almost all other malarious regions of the world (D'Alessandro and Buttiens 2001; Arrow, Panosian et al. 2004).

1.2.3 Proliferation of research and decision making bodies

In the later decades of the 20th century, the persistent burden of malaria gradually re-ignited research into alternative interventions for diagnosis, treatment, and vector control. This was accompanied by the emergence of multilateral national and non-governmental bodies merging malaria related research with policy making (Alilio, Bygbjerg et al. 2004). In addition to product research and development, bodies such as the *Multilateral Initiative on Malaria* (MIM) and the WHO *Tropical Disease Research programme* (TDR) and later *Roll Back Malaria* (RBM) partnership, encouraged the education and training of health researchers, professionals and policy makers in SSA and elsewhere in the developing world, thus further empowering decision making bodies across the continent.

The major challenge facing these entities was the demise of CQ as an effective first line treatment, requiring the introduction of alternative therapies. Unfortunately, none of the

alternate options combined all the advantages of CQ. Perhaps due to a reluctance to part with CQ, the process of consideration and adoption of new antimalarials was slow. In Kenya it took over 14 years from the initial evidence documenting CQ resistance to the final policy change (Shretta, Omumbo et al. 2000). In Tanzania the first line treatment had to be changed twice in the space of 5 years as by the time the first change was formally adopted, the new drug was already facing high treatment failure rates (Williams, Durrheim et al. 2004; Mubyazi and Gonzalez-Block 2005). The introduction of new first line drugs and the rapid development of resistance to these became a menacing feature of malaria control in SSA and Southeast Asia in particular (Bloland, Kachur et al. 2003).

Despite these setbacks, the last decade has seen a vast increase in global activities relating to malaria control – specifically in the proliferation of malaria-based agencies, and unprecedented funding for these. Bates and Herrington (2007) provide a fairly optimistic overview of malaria control in recent years, describing how the decision making process has been transformed from one dominated by centralized bodies and special interests to one that is more transparent, inclusive, and reactive to local circumstances.

Despite the proliferation of decision making bodies, new interventions and strategies often continue to be promoted on a continent wide scale, even though their use may not be appropriate in all circumstances. This has been cautioned against in the case of rapid diagnostic tests (Bell et al., 2002), artemisinin combination therapy (Whitty and Staedke, 2005) and home management of malaria programmes (Hopkins, Talisuna et al., 2007). The shifting of decision making powers to local bodies could ensure that only those interventions that are best suited to local circumstances are adopted.

1.3   The introduction of artemisinin combination therapies and rapid diagnostic tests

Other developments have also provided further promise of improved malaria control. Since the late 1990s, a new class of drugs has been introduced with the promise of filling the gap left by the demise of CQ as an effective antimalarial. Artemisinin compounds,

24

identified in Chinese herbal medicine as an effective treatment for fever, have been developed for use as antimalarials and have proved highly effective in rapid elimination of infection. To further ensure therapeutic success and reduce the probability of resistance emerging to them, artemisinin compounds are being used in combination with other antimalarials.

Presently, artemisinin combination therapies (ACTs) hold much promise for effective treatment of malaria as has been demonstrated in much of Southeast Asia (Nosten and Ashley 2004) and more recently in a number of locations in Africa (Barnes, Durrheim et al. 2005; Bhattarai, Ali et al. 2007). ACTs are now being heavily promoted for first line use in the treatment of uncomplicated malaria across SSA by the WHO and other influential bodies such as the Global Fund to Fight AIDS, Tuberculosis and Malaria, and the US Institute of Medicine (WHO 2001; Arrow, Panosian et al. 2004; WHO 2006b).

A host of arguments, however, have been made that ACTs should not be given as presumptive treatment, and that countries should rely more heavily on parasitological diagnosis prior to their use (Bloland, Kachur et al. 2003; Amexo, Tolhurst et al. 2004; Breman, Alilio et al. 2004; Nosten and Ashley 2004; Whitty and Staedke 2005). Three main areas of concern have been raised. First, these therapies come at a cost significantly higher than any of their predecessors, raising concerns around the sustainability of using them as first line treatment. Second, there is the possibility of encouraging resistance to artemisinin compounds, with very limited options for their substitution in the near future (Bloland and Kachur 2003). Third there are safety concerns regarding their widespread use, particularly amongst young children and pregnant women (Price 1999). The increased focus on improved diagnostics has also emphasised the potential benefit of demonstrating when fevers are not caused by malaria parasites - allowing clinicians to consider and treat alternative causes of illness (Shillcutt, Morel et al. 2008).

The arguments in favour of improving diagnostic practices have coincided with the increasing availability of rapid diagnostic tests (RDTs), which are competing with the traditional use of microscopy for the detection of parasitaemia. RDTs allow for fast and

relatively accurate diagnosis and require only minimal training and infrastructure for their use (Moody 2002).

With an increased demand for ACTs and RDTs, and a vast rise in available funding from donor organizations (Waddington, Martin et al. 2005), a growing variety of ACTs and RDTs are becoming available on the market. This increase in antimalarials and diagnostic tools is accompanied by a host of evaluations considering their efficacy, effectiveness and costs in a variety of settings. The consequence is a relatively large body of data accumulating rapidly on ACTs and RDTs. These data, however, are not amenable for use by decision makers due to their fragmentation in three main areas: (1) the specific antimalarial or RDT; (2) the target sub-groups in the population; (3) the geographical location.

In terms of the particular intervention, the variety of available tests and proposed drug combinations in the ACT class hampers efforts to agree in broad terms on their efficiency. This was recently demonstrated in a meta-analysis of artemether-lumefantrine in a four dose regimen which concluded that this was less effective than the non-artemisinin antimalarials it was compared to (Omari, Gamble et al. 2004). This was followed shortly by a meta-analysis of the six dose regimen of the very same combination that came to the opposite conclusion (Omari 2005). Similarly, while overall RDTs have been shown to be of sufficiently high accuracy to be considered for use in routine practice (Murray, Gasser et al. 2008), individual studies of RDTs detecting the same antigens and even of the same brand have reported discrepant results (Tjitra, Suprianto et al. 2001; Guthmann, Ruiz et al. 2002; Swarthout, Counihan et al. 2007).

Regarding the target population, there are a number of factors relating to patient profiles that will influence the interventions' effectiveness. The patient's age for instance might determine how susceptible they are to infection and subsequent development of disease due to acquired immunity (Greenwood, Bojang et al. 2005). This will influence both the efficiency of the treatment used, and also the utility of using diagnostic tests to confirm the existence of malaria parasites.

Other patient characteristics that might influence efficiency are HIV status and pregnancy. The interaction between HIV and malaria is garnering attention and is likely to influence the efficiency of both diagnostics and treatment (Kublin and Steketee 2006; Laufer, van Oosterhout et al. 2006; Van Geertruyden, Mulenga et al. 2006a; Van Geertruyden. Mulenga et al. 2006b). Pregnant women can lose their immunity to malaria, therefore fast and accurate demonstration of parasitaemia is critical. Furthermore there have been concerns regarding the safety of certain antimalarials for this subgroup, therefore presumptively treating them where prevalence is low may be inappropriate (Breman 2001; Brentlinger, Behrens et al. 2006).

Regional differences such as those in transmission patterns and resistance to partner drugs will also influence intervention effectiveness (Francis, Nsobya et al. 2006). ACTs will be less effective when the partner drug in the combination is one that is locally failing in mono-therapy (Whitty and Staedke 2005). Presumptively treating all febrile patients where transmission is very low is also unlikely to be efficient. Competing interventions and health needs, as well as the available health budget, will all impact on whether the use of specific malaria diagnostics and treatments are an efficient use of resources. These are all factors that vary widely by location.

Many of these factors are highly temporally dynamic, adding a further complication. Treatment efficacies for instance vary, as do transmission intensities and host immunity in the population. These variations are due to factors such as emerging resistance, urbanization and climate change. Given these circumstances it is clear that no single choice of drug, diagnostic, or case-management strategy will be appropriate for all patients, at all times and all locations. An expensive ACT may not be the most appropriate first line drug in all settings, for instance, in the treatment of semi-immune adults in areas of low resistance to cheaper drugs. On the other hand its use in presumptively treating certain subgroups, such as young children in high transmission areas, may be beneficial.

There is a sense of urgency to these issues, both to curb the ongoing toll of malaria and other non-malarial febrile illnesses (NMFIs) in the most vulnerable populations in SSA, but also to temper the drive for the immediate widespread deployment of ACT and RDTs

across SSA. Health care systems already buckling under the strain of illnesses such as malaria cannot afford to pursue ill-suited strategies to tackle them.

As this section has discussed, malaria research and policy making has evolved dramatically over the past century from the days when a single body advocated a small number of uniform interventions across the continent. Today policy makers face a mixed blessing of numerous options for diagnosis and treatment of malaria, with an abundance of data on the interventions under consideration, and on a host of other factors that should influence their decisions. With more choice comes more decisions for which, as will be discussed in later chapters, many standard economic evaluations fail to provide adequate answers or recommendations.

So far this chapter has provided an overview of malaria, research and policy. The next section will provide an overview of malaria pathology and epidemiology, and the impact malaria has on the economy as a whole. This is followed by a description of the different approaches to incorporating research with decision making in formulating policy, such as choice of malaria diagnostics and treatment strategies.

1.4 *Overview of malaria pathology*

1.4.1 The malaria life cycle and interventions for its control

While establishing a precise definition of clinical malaria has proved challenging, the initial cause of illness is straightforward – an infection with a protozoan parasite from the genus *Plasmodium* (Greenwood, Bojang et al. 2005). Almost all human infections are caused by four species of the parasite, with *P.falciparum* causing the greatest health burden. The life cycle of the parasite is such that it evolves through a number of stages in both its human and mosquito hosts, as depicted in Figure 1-1.

THE LIFE CYCLE OF MALARIA

MOSQUITO ⟷ HUMAN

Sporozoites migrate to salivary gland

Mature oocyst bursts and replaces sporozoites

Oocyst grows

Sporogony

Ookinete attaches to gut and becomes oocyst

Fertilization

Fertilized female gamete forms a zygote which becomes an ookinete

Mosquito injects parasites

Hepatic schizogony

Erythrocytic schizogony

Female mosquito removes parasites

Sporozoites migrate to the liver

Merozoites leave the liver and infect red cells

Trophozoites mature in red cells to become either schizonts or gametocytes

Schizonts release more merozoites

Gametocytes wait for mosquito

Figure 1-1: The life cycle of malaria. Taken from http://www.emro.who.int/RBM/AboutMalaria-QuickOverview.htm - Accessed 15/11/07

From the public health perspective the different stages of the malaria life cycle are targeted by different interventions, aiming at either clinical care or at reduction of transmission between humans and vectors (Shiff 2002).

After infection from the bite of a female *anopheline* mosquito, a number of sporozoites migrate to the liver where they remain for an average period of 6.5 days. The initial liver stage sees the reproduction of the parasite; at this point the host is asymptomatic (therefore also unaffected by treatment). The only public health intervention relevant to this stage is the development of malaria vaccines that would terminate the cycle at this point (Moorthy, Good et al. 2004).

At the next stage the parasite develops into merozoites that enter the bloodstream, and release tens of thousands of merozoites, first invading red blood cells and then attaching themselves to walls of small blood vessels, potentially blocking the blood flow to organs. This process repeats itself cyclically with parasite loads increasing exponentially. At this stage the host can become symptomatic at varying degrees of severity, ranging from mild fever to severe illness and death. Antimalarial guidelines target rapid provision of the drugs to halt this cycle and reduce the parasite load. Different drugs perform this action at

varying speeds and effectiveness, while some will have a further advantage of a limited prophylactic effect (White 2005). More persistent drugs, however, also expose themselves to a greater risk of development of parasite resistance as they drop below therapeutic levels.

Undisturbed, the parasites would enter the next stage where they are differentiated into male and female gametocytes. Although benign to the host, these are the forms responsible for transmission back to mosquitoes. From a public health perspective, therefore, an antimalarial that reduces gametocyte load could also contribute to reduced transmission. This is a particular characteristic of artemisinin compounds, which suggests they might have the additional benefit of reducing overall transmission (Obua, Okell et al. 2008).

A female mosquito that is infected with gametocytes during feeding will then carry the parasite as it evolves back to the sporozoite stage for re-infection of another host. The main interventions aimed at reducing transmission are targeted at vector control, either by reducing the number of mosquitoes altogether, as in the widespread spraying of insecticides and with the elimination of breeding grounds, or by protection of susceptible humans through indoor residual spraying or the use of insecticide treated bed-nets. A range of other interventions are in development that also aim at eliminating transmission, perhaps most notably genetically modified mosquitoes that are fully refractory to *Plasmodium* parasites (Greenwood, Bojang et al. 2005).

1.4.2 Clinical manifestations

Infection in the human host can manifest itself in a number of ways. Most well known are the intermittent febrile episodes termed *uncomplicated malaria*. A typical clinical episode will consist of a series of febrile episodes of varying length and frequency, depending in part on the species and on host susceptibility (Arrow, Panosian et al. 2004). Other accompanying symptoms can include head and body pains, cough and diarrhoea (Snow and Marsh 1998).

More acute infections can induce seizures and coma, referred to as *cerebral malaria,* with estimated case fatality rates ranging between 10%-40% (Arrow, Panosian et al. 2004). Another manifestation, usually associated with repeated infections, is *severe anaemia,* also a life threatening condition. The third and increasingly recognised manifestation of severe malaria amongst African children is *respiratory distress and metabolic acidosis (English, Waruiru et al. 1996).* While not the most common form of severe malaria, this is the most fatal, particularly when coinciding with other manifestations of the infection (Greenwood, Bojang et al. 2005).

1.4.3 Long term morbidity

In addition to the above symptoms directly associated with the infection, a number of long-term conditions are possible. Higher prevalence of anaemia is closely correlated with infection rates, also among non-febrile individuals, as is a compromised immune system and higher susceptibility to other diseases (Winstanley, Ward et al. 2004). Long term neurological impairment can also follow cerebral malaria (Arrow, Panosian et al. 2004).

Pregnant women are particularly susceptible to infection, with higher risk of adverse outcomes for both themselves and their infants, expressed in anaemia amongst mothers, and in higher probability of stillbirths for infants. Parasitaemia amongst pregnant women is also associated with low birth weight, itself linked with risk of neo-natal mortality and morbidity later in life (Molineaux, Muir et al. 1988).

1.4.4 Host immunity

The cause of malaria illness will by definition be malaria parasites, though in endemic areas, where the vast majority of infections will not result in clinical manifestations, the prevalence of parasitaemia amongst asymptomatic individuals in high transmission areas can be as high as 90% (Bottius, Guanzirolli et al. 1996). Older children and adults in these areas will develop immunity to the parasites, often showing no clinical manifestation beyond benign mild febrile episodes, indistinguishable from fever resulting from other causes.

The host's degree of susceptibility to infection and development of severe disease is largely a product of immunity to parasites due to prior infections, thus age and transmission intensity will play a role in determining the outcome of infection. While these factors go part way in explaining the diversity in outcome, much remains to be understood as infected hosts of similar characteristics will often experience remarkably different outcomes (Arrow, Panosian et al. 2004). Broadly however, immunity to severe manifestations and death is acquired early on in life while immunity to other less severe manifestations is more gradual (Gupta, Snow et al. 1999).

1.5 *Case definition and diagnosis of malaria*

In non-endemic areas the definition for clinical malaria is relatively simple – the presence of *Plasmodium* parasites accompanied by a history of fever (Marsh and Snow 1999). This, however, is insufficient in high transmission areas where a large proportion of the population can be parasitaemic and asymptomatic at any given time. Attempts to resolve this by using parasite load cut-off points have proven controversial as parasite densities do not always correlate directly with clinical incidence or with disease severity. A more recent method has somewhat improved this with the use of logistic regression to estimate the risk of fever as a continuous function of parasite density (Mwangi, Ross et al. 2005). While this can provide more accurate measures of the overall burden of malaria in the total population, it is still an inaccurate tool for determining whether a particular patient's cause of illness is their infection with *Plasmodium* parasites.

The ambiguity in case definition and diagnosis is further exacerbated by the inability in many instances to obtain a diagnosis of parasitaemia due to limited infrastructure, with clinicians relying entirely on non-specific clinical symptoms for treatment decisions. Attempts to construct algorithms to improve the specificity of clinical diagnosis of malaria have so far proved unsuccessful (Chandramohan, Jaffar et al. 2002).

Given these difficulties, when CQ was still effective and resistance was not yet a looming threat, presumptive treatment for all age groups and in all endemic areas was a rational policy, particularly given CQ's antipyretic qualities relieving symptoms of other illnesses.

This however, may well have contributed to the confusion of malaria with other illnesses on the part of patients and clinicians, as patients given an antimalarial felt their symptoms alleviated. The confounding of malaria and febrile illness is not limited to patients and clinicians in remote areas where diagnostics are inaccessible, but permeates the world of malaria research and control in assessments of disease burden and the implementation of malaria control measures.

1.6 *Consequences of overdiagnosis of malaria*

Febrile illnesses treated as malaria account for approximately 20%-40% of outpatient visits in SSA (Chima, Goodman et al. 2003). There is a shortage of data concerning the exact break down into malaria and other NMFIs. There are, however, a number of studies in SSA suggesting that malaria is being grossly over diagnosed, especially in low transmission areas (Biritwum, Welbeck et al. 2000; Makani, Matuja et al. 2003; Reyburn, Mbakilwa et al. 2007). This raises concerns regarding detrimental consequences for patients suffering from NMFIs, where the case fatality rate can be higher than that for malaria, particularly when receiving antimalarials instead of appropriate treatment (Makani, Matuja et al. 2003).

The overdiagnosis of malaria has ramifications beyond the patient at hand. Notwithstanding the potential benefit of treating a febrile patient with an antimalarial, there are societal costs associated with each treatment given, described here as the "harm of treatment". This potential harm associated with the provision of an antimalarial or antibiotic includes the potential for drug toxicity (Price 1999), the contribution to the development of parasite (or bacterial) resistance (Bloland, Kachur et al. 2003), and the opportunity cost of the use of scarce resources.

From a public health perspective, the confusion surrounding case definition and diagnosis also results in considerable variation in total estimates of morbidity and mortality, with the speculated number of clinical episodes of malaria ranging from 400 million to 5 billion clinical episodes annually, leading to anything between 1 and 3 million deaths a year (Breman 2001). The magnitude of even the lowest of these estimates justifies placing malaria high on health agendas; however this uncertainty can also lead to pursuit of highly

inappropriate strategies, where billions of clinical episodes might be incorrectly treated and up to two million lives might be lost every year after incorrect diagnosis with malaria.

Assessments of the true burden of malaria are continuously challenged by factors such as urbanization (Roberts, Macintyre et al. 2003), climate change (Tanser, Sharp et al., 2003), and the widespread implementation of interventions such as insecticide treated bed-nets (Curtis, Jana-Kara and Maxwell, 2003). These all affect transmission intensity and consequently prevalence of malaria and host immunity, posing substantial difficulties to policy makers considering treatment and diagnostic strategies, and raises concerns as to how they might account for such variation in their decision making process.

1.7   The impact of malaria on the economy

In addition to the direct health burden malaria places on populations at risk, there are further ramifications that can be detected across the entire economy. A number of studies have been carried out, exploring the impact malaria has on the economy at both the micro and macro levels. These studies express a range of views on how malaria affects the general economy, and the possible outcomes for strategies that might reduce or eliminate the presence of malaria.

One of the earliest evaluations to take a macro-economic approach is an analysis of the impact of malaria eradication in Sri Lanka on average income per capita (Barlow 1968). Following Barlow's estimate of the increase in child survival rates, the long term projection was that population growth would lead to a downturn in income per capita, suggesting that from a purely economic perspective, malaria eradication would result in poorer economic performance.

Two more recent studies have assessed the overall economic burden of malaria by assessing its impact on growth rates. Gallup and Sachs (2001) carried out a regression analysis for the correlation between malaria and economic growth. Their results suggest that countries with intensive malaria grew 1.3% less per person per year, and a reduction

of 10% in malaria was associated with 0.3% higher growth (Gallup and Sachs 2001). McCarthy, Wolf and Wu (2000) also carried out a regression analysis but allowed for two way causality between morbidity and economic growth rates, resulting in a lower estimate of 0.25% reduction and far greater variability in results (McMarthy, Wolf et al. 2000).

Another approach to determining the impact of malaria on the economy is through micro-analyses that sum up total household or government expenditure on treatment and productivity losses related to malaria morbidity and mortality. Shepard et al (1991) used such an approach based on costing from four countries to extrapolate to the total burden of malaria in Africa, estimating the total cost at 1.1 billion USD (Shepard, Ettling et al. 1991). Chima, Goodman and Mills (2003) later summarised the findings of such analyses, demonstrating their shortcomings in adequately evaluating the true costs of malaria to the economy, notably the overestimation of the productivity losses associated with uncomplicated malaria, and the underestimation of the economic burden of severe malaria (Chima, Goodman et al. 2003).

Factors such as the impact of malaria on long term educational attainment and subsequent productivity have yet to be thoroughly explored, although the overall negative correlation between malaria and average schooling years has been demonstrated, and can be assumed to have a substantial negative impact on the economy as a whole (Lucas 2006).

1.8 Approaches to decision making

The high profile of malaria on the international agenda, and its overdiagnosis in clinical care and burden of disease assessments, are one potential source of bias in the decision making process concerning the adoption of new malaria control interventions. Another potentially biasing factor is the influence of the decades of presumptive treatment with CQ, being affordable, accessible, and effective for both malaria and other febrile illnesses. This may well have resulted in a temptation to fill the void left by CQ by distributing newer antimalarials in a similar fashion. Such biases sway policy makers when their decisions are made on a purely intuitive basis. Use of analytic decision models can ensure that these

factors are accounted for in a more rational manner. In this section a brief overview is given on how policy makers could draw on research to inform their decision-making in the face of the pressing and complex issues they are expected to act on.

Deciding on policy such as switching first line drugs for malaria involves the consideration of a multitude of factors, many of which exhibit high levels of uncertainty. Other inputs might be of a more value laden nature; for instance, where policy makers might have to trade off future costs and benefits for present ones, often at an unequal weight. Under such circumstances of uncertainty and the presence of value laden factors, a number of distinct approaches to decision and policy making can be found.

1.8.1 The status quo bias

When having to choose between competing strategies, the simplest reaction would be to maintain the *status quo*, a common response when facing conditions of change and uncertainty (Samuelson and Zeckhauser 1988). In fact the '*status quo* bias' has been estimated to increase with the complexity of a situation and the number of proposed alternatives. The ongoing harmful consequences of the *status quo* can continue to remain low on the political agenda when both policy makers and the public have already resigned themselves to this reality. Only when the issues reach some threshold of unacceptability and recapture public attention is the decision maker shaken into action, albeit later than might have been warranted.

This pattern was evident in the reluctance to switch to the use of new antimalarials even when the existing ones were increasingly recognised to be ineffective, as described by Shretta et al. (2000), examining the policy change from CQ to SP in Kenya, in relation to the evidence on CQ resistance. The authors identified studies spanning a period of 14 years prior to the final change from CQ, which demonstrated the fall in its efficacy. However, only when the evidence became irrefutable with a number of studies with failure rates as high as 72% and adverse clinical outcomes, were policy makers more proactive in their decision making.

1.8.2 Intuitive decision making

The dominant form of decision making in private life is an intuitive one, which simply relies on our capacity to digest numerous sources of information relating the probabilities and outcomes of different events occurring, and subsequently react in the most adequate manner (Hammond 2000). Although a highly complex process, this is how everyday decisions are routinely made. Facts, uncertainties and values are all constantly merged together, often in a manner that even the individual decision maker cannot tease apart. A parent to a febrile child, for instance, will intuitively merge their subjective assessment of the probability of this being malaria or a different illness, and the probability of a severe outcome, along with factors such as the different costs associated with the possible responses (buy medication at the drug store, spend a day at the outpatient clinic etc.), and act accordingly.

While this form of decision making might suffice in private life, it has a number of limitations when facing broad policy considerations in the public sphere. First, most people have a fairly limited capacity to analyze intuitively even mildly uncertain situations, particularly when these involve a number of factors, often with conditional probability of occurrence. This has been repeatedly demonstrated in the literature relating to heuristics and decision analysis (Tversky and Khanerman 1974). The misinterpretation of diagnostic test results, for example, is a recurrent problem in clinical practice, when clinicians all too often confuse test sensitivity with its positive predictive value[1] (Altman and Bland 1994). This is but one example amongst many that are revealed when decision making processes are mapped out systematically (Hammond et al., 1997).

Second, given these potentially flawed decision making processes, these are often accompanied by a complete lack of transparency, as it can be unclear what information

---

[1] A positive test that is 97% sensitive does not mean that there is a 97% chance that the patient is carrying the condition of interest. The latter probability, or predictive value, is an amalgamation of test sensitivity, specificity, and the prevalence of the condition amongst the population.

entered the decision making process and the value attributed to different inputs. This can impede the identification of influential but uncertain parameters that might require better estimation in future decision making. The lack of transparency also implies less accountability for bad decisions, absolving the decision maker from scrutiny in how decisions were reached, and avoiding the need to invest appropriate resources in considering the different options. Despite these limitations this form of decision making has been acceptable and even revered in both clinical medicine and policy, and only in recent years has it made way to slightly more analytical approaches (Davies and Nutley 1999).

1.8.3 'TIABIM' (Taking into account and bearing in mind)

In recent years it has become the norm in both clinical medicine and policy to adopt, or at least pay lip-service to, more evidence-based approaches to decision making. The TIABIM approach (Dowie 2005) is often carried out by compiling some of the relevant evidence systematically so that parameters such as clinical efficacy and costs are fed to decision makers, to be subjected to an intuitive synergy with a range of other factors.

The proliferation of 'Health Impact Assessment' reports is indicative of the will to incorporate evidence on a range of factors in the decision making process (Dowie 2003). While this represents some progress from entirely intuition based decision making practices, it still places unfair and often impossible demands on policy makers to synthesise all evidence for and against the intervention in what remains a highly intuitive manner. As will later be described, many economic evaluations opt for a very narrow scope of factors to be included in the analysis, such as immediate intervention costs and intermediate measures of outcome. Antimalarials for instance, are often evaluated based on their parasite clearance rate in a particular place and time, leaving the onus on the decision maker to intuitively merge this with other factors such as the possibility of emerging resistance, and extrapolate the significance of these for the long term or to other settings.

As with the intuitive decision making approach described earlier, the drawbacks of this approach relate both to the process itself, and the ability to scrutinize it for improvement. In order to improve both the quality of the decisions and their transparency, a more analytical approach is required.

1.8.4 Decision analysis

According to the Society of Medical Decision Making, decision analysis (DA) can be defined as "a methodology for making decisions by identifying alternatives and assessing them with regard to both the likelihood of possible outcomes and the costs and benefits of the outcomes."[2]

DA facilitates the handling of situations where a decision is required despite high levels of uncertainty concerning factors that determine final outcomes, which denies the decision maker an obvious choice of action. In its simplest form, decision analysis requires the use of a model representing the alternative courses of action, the identification of all possible outcomes related to these options, and the assignment of probabilities and payoffs to each of these. Once relevant parameters have been identified and placed in the model, an expression of the uncertainty surrounding these is required in the form of the probability of different values for the parameter. The analysis then requires that these be complemented by the utilities for the different outcomes, as defined by the decision makers (Lindley 2000). The option with the highest expected payoff is the one recommended for adoption.

In addition to this primary function, such models allow further exploration of how results are affected by individual parameter uncertainty. This can be beneficial for both policy and research purposes in providing indication as to which parameters are most influential for the final outcome, and which would benefit most from further research.

---

[2] http://www.smdm.org   Accessed July 18th 2007

While the use of DA represents an improvement in systematic and rigorous policy making over intuitive decision making, paradoxically it can at times appear as insufficiently scientific from the researcher's perspective. For instance, one of the defining characteristics of the DA approach is that it requires a departure from a classical scientific method, where a parameter estimate will be accepted as true if the null-hypothesis is disproved following the finding of a p value below 0.05. In DA, the luxury of maintaining these standards for determining whether a hypothesis is true or false cannot be afforded, as decisions will still be required even in the absence of conclusive evidence. Instead, the probability of the truth of the hypothesis is introduced into the analysis, allowing this uncertainty to carry through to final decision recommendations (Dowie 2005).

DA is characterised by a decision oriented approach, in that its starting point is structuring the problem in an analytic manner, usually a decision model, and then seeking out the best available evidence to populate it. Consequently, the use of decision models has been greeted with a number of criticisms. These include the possible use of clinical trials with insufficient statistical power to inform parameter values, the use of observational studies with potential biases, and the validity and transparency of the models (Buxton, Drummond et al. 1997). There is an underlying concern that too many of these factors are left to the discretion of the modeller. At the extreme, the lesser degree of scientific rigour can lead to manipulation of decision recommendations and potentially the abuse of the methods to promote commercial agendas (Kassirer and Angell 1994).

While these are all valid concerns, the use of decision models must be compared to the prevailing decision making processes, characterised by far less systematic approaches, where the evidence is at best partly digested by the decision maker who will then having 'borne everything in mind', conclude that one course of action is preferable to another. The existence of biases and commercial interests in such processes is also far less visible.

1.9 *Decision analysis and economic evaluations*

Decision analysis is a general approach to problem solving. While there are a variety of types of economic evaluation frameworks (see Chapter 3), most function by applying the

same approach as that of DA – identifying the alternative options involved, estimating the probabilities of different outcomes for each choice of action, and assigning costs and benefits to these outcomes.

How informative economic evaluations are for policy making purposes will depend on their framework, content, and how they handle variability and uncertainty in parameter estimates. Economic evaluations often limit themselves to the inclusion of immediate costs and intermediate measures of outcome, providing decision makers with measures such as the 'cost per patient correctly diagnosed', that they will have difficulty in interpreting in relation to other considerations (Drummond, O'Brien et al., 2005). Other factors that have significant bearing on the effectiveness of an intervention such as patient and/or clinician adherence to its use are often excluded from analyses, which instead focus only on the efficacy in trial circumstances. Evaluations could alternatively seek greater decision relevance by trying to incorporate longer term costs and benefits and ensure their relevance to the particular intervention and population being considered.

The general theme of this thesis is that economic evaluations should as far as possible seek to maximize their relevance to decision making context. Doing so requires that all major factors are included in the scope of the analysis, rather than being left for intuitive integration, regardless of the difficulties in estimating these. Another requirement is that evaluations be based on data applicable to the location and population of interest. Practical considerations however usually imply that data collection is limited to a small number of sites, with an inevitable delay between data collection and its publication, to allow for meticulous data collection methods, analysis, and the existence of a safeguard in the form of a peer review process (Buxton, Drummond et al. 1997). While parameter estimates derived from this process can claim internal validity, evaluations based on such estimates will potentially have limited generalisability of results and may be based on outdated data. This is of particular significance in the context of malaria where many of the factors that inform an evaluation are highly dynamic, both temporally and geographically. Some of the most recently published evaluations of ACTs and RDTs, for instance, are based on data from 2005, when ACTs were over twice the cost at the time of publication

41

(Rolland, Checchi et al. 2006; Lubell, Reyburn et al. 2007), and these costs are expected to continue to fall in the near future.

This is but one in a long list of variable factors, that also includes treatment effectiveness, transmission intensities and RDT accuracy (due to the development of new brands). Consequently policy makers consulting the most recently published evaluations are likely to be misinformed by evaluations that were relevant in a different time and place.

1.10 *Chapter conclusion*

This chapter has given an overview of the history of malaria research and policy. It further outlines the different modes of decision making and how they shape policy. Devising malaria diagnostic and treatment strategies, policy makers can follow different approaches to decision making. The more intuitive practices that characterise decision making in health care as a whole are particularly vulnerable to bias, especially in the context of malaria, towards overdiagnosis at the expense of other illnesses. Other barriers to effective decision making are the wide range of factors beyond immediate costs and benefits that are associated with malaria control, and the significant variation in the manifestation of malaria in different geographical areas and over time.

As was described, an increasing number of decision making bodies focused on malaria control are proliferating in much of the developing world. This has the potential to ensure that only those interventions that are most suited to local circumstances are indeed adopted. This contrasts sharply with the past decision making structures which can be characterised as monolithic and unresponsive to local variation.

This thesis sets out to ensure that decision making tools are available to enhance this potential further, by making use of some of the most recent methods used in the developed world for the evaluation of new health care interventions, to ensure that decision models are comprehensive in the range of input factors they account for, and that the output they provide is responsive to the different circumstances pervade in the areas where the interventions are considered for use.
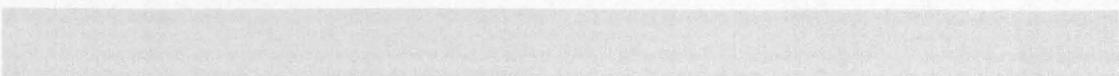
2. Review of the literature on economic evaluation of malaria diagnostics and treatments

2.1 Aims of the review

The aims of this chapter are to review the literature concerned with economic evaluations of malaria treatment and diagnosis, and discuss the modelling approaches they have employed, with a focus on their value for decision making. This is concluded by outlining the gaps and weaknesses in the existing literature that this thesis addresses in subsequent analyses. Annex 1 provides details of the methods and results of each of the studies.

2.2 Review scope and structure

There is a wide range of methodologies for conducting economic evaluations which have been categorized and described below. The category order, however, does not indicate their merit or relevance to decision making requirements, but does reflect how comprehensive the analyses are, ranging from those that included only immediate trial results for the patients in particular, to those that model the long term impacts of the interventions for society in general. The main focus of the review is on the methods used in the studies and how these can be enhanced in future evaluations. The actual study results are less relevant and are summarised briefly in Annex 1.

The review opens with trial based evaluations that did not make extensive use of modelling techniques. These evaluations rely almost entirely on patient level data on cost and effectiveness data from a single trial to draw conclusions regarding the comparative efficiency of the interventions under investigation. Following these are those evaluations that compile data from a range of secondary sources to estimate the costs and

effectiveness of interventions of interest, without making use of a formal decision model structure, but rather calculate average and incremental cost-effectiveness ratios.

The third category consists of decision tree based evaluations. Numerous analyses made use of this structure, although it should be noted that this does not necessarily imply that a decision analytic approach has been adopted. Some analyses have made a limited use of decision trees simply for breaking down and portraying complex problems[3]. The fourth category consists of system dynamic models. The main defining feature for these models is their treatment of disease transmission and the risk of infection as endogenous to the model, making use of mathematical tools such as differential equations to define the relationship between model variables.

This order of categories represents to a certain extent the model complexity, therefore where an analysis used more than one of these methods it was placed in the higher category. Each of the following sections provides a brief introduction to the general approach, and a discussion of the methods that were applied in the evaluations.

2.3 Search strategy

Search strategy: The literature review was completed in February 2008 and aimed to identify all relevant published economic evaluations of malaria treatment and diagnostics to date. A number of non-published studies were also identified by consultation with experts in the field.

---

[3] The terminology and categorization varies in the literature, and the term 'decision modelling' is often used to encompass all evaluations based on a range of secondary sources; a distinction is made here between evaluations that limit their aim to providing a measure such as an incremental cost-effectiveness ratio, and those that provide actual decision recommendations, for instance by incorporating the policy makers' willingness to pay for the benefits provided.

2.3.1 Inclusion criteria

Intervention:

- Use of antimalarial drugs for the treatment of *P.falciparum* malaria, being the cause of most of the total health burden of malaria due to its higher mortality rate than the other *Plasmodium* species

- Use of diagnostic tests to determine parasitaemia

- Implementation of different strategies for the management of malaria suspected patients

Study criteria for inclusion:

- Economic evaluations based on RCT results

- Evaluations based on the modelling of secondary data

- Location – Low and middle income countries

Databases searched: Pubmed, Cochrane, Embase, ID21, HEED, WoS. These were initially searched in April 2006 and results were updated in February 2008.

MeSH terms: Economic evaluation, Cost/cost analysis, malaria/Drug therapy, diagnosis, antimalarial

Text search terms: malaria, plasmodium falciparum, febrile illness, economic evaluation/analysis, artemisinin, rapid diagnostic test, RDT, Immunochormatic tests, microscopy, presumptive treatment.

Where text searches were used these were done by searching for each phrase individually, then combining search results.

References from the relevant studies were used to identify any further papers. Results were also compared with researchers working on related topics.

2.4 Search results

Twenty-two economic evaluations of antimalarial drugs and diagnostics were identified in total. Three additional studies relating to costs or outcomes, that did not qualify as economic evaluations, are also presented as they introduced noteworthy methodologies.

2.4.1 Category I - Trial based evaluations

Seven studies used primary data from single trials to determine cost-effectiveness of malaria treatment, diagnostics and case-management strategies (Jonkman, Chibwe et al. 1995; Honrado, Fungladda et al. 1999; Bualombai, Prajakwong et al. 2003; Gogtay, Kadam et al. 2003; Fernando, Karunaweera et al. 2004; Wiseman, Kim et al. 2006; Chanda, Masiye et al. 2007). There is an apparent advantage to such evaluations in that they are widely accepted as robust given their high internal validity, drawing conclusions directly from a trial with a design that is supposed to ensure a statistically significant result.

Nonetheless three main areas of concern in trial based economic evaluations may be gleaned from these studies: (1) the adequacy of the effectiveness measures; (2) the generalisability of results to other settings; and (3) the interpretations of resulting cost-effectiveness ratios. In addition there is the special case of multi-centre studies, with the challenges they pose to interpretation of variability between study sites (Chanda, Masiye et al. 2007).

Determining intervention effectiveness on the basis of clinical trial results can be problematic for four reasons. First, clinical trials often assess treatment efficacy rather than effectiveness. Thus they might not represent routine clinical practice in terms of the patients being treated, who are often selected by stringent criteria, or the care they are given, which will be protocol driven rather than that being delivered in the routine provision of care (Soto 2002).

Second, due to budget and time constraints, clinical trials will often have a shorter than ideal period of follow-up, relying on intermediate outcomes to determine the impact of interventions which may in fact not reflect final health outcome (Brennan and Akehurst 2000). In the case of antimalarials, duration of follow up is often shorter than ideal due to

46

the likelihood of reinfection, which without the appropriate technology, cannot be distinguished from recrudescence.

Third, choice of the 'gold standard' against which effectiveness is measured in the trial could have undue impact on results. In both Bualombai et al. (2003) and Fernando et al. (2004), RDT accuracy was measured against the performance of microscopy, both as comparator for routine diagnosis, as well as functioning as the gold standard. Use of the same technology as comparator and gold standard is likely to bias results against RDTs, as was demonstrated by Bell et al. (2005) in a re-evaluation of a previous trial comparing diagnostic accuracy for RDTs and microscopy, but including PCR as the gold standard.

Fourth, the generalisability of results to other settings is of much concern when cost-effectiveness estimates are drawn from a single setting (Sculpher, Pang et al. 2004). While authors often express the need for further studies in other locations, results are often used to provide more general policy recommendations. Honrado et al. (1999) for instance, concluded on the basis of their trial, which included a total of 137 patients in a single setting, that artesunate is more cost-effective and should be recommended for use in treating uncomplicated Plasmodium falciparum malaria in clinics across Thailand. Similarly Gogtay et al. (2007) saw their analysis as sufficient to inform policy makers regarding the cost-effectiveness of the different treatments "at least for referral centres like ours" (p.879).

A variety of factors impede the ability to generalise trial results to other settings. In the context of malaria, of most significance are transmission intensity, host immunity, and parasite resistance to antimalarial drugs (Snow 2000; Whitty and Staedke 2005). Transmission intensity will be of particular significance under a strategy of presumptive treatment due to the varying proportion of patients that benefit from treatment. In a recent study in Tanzanian highlands, it was found that while parasitaemia rates were under 1.5% amongst febrile patients, over half of these still received an antimalarial (Reyburn, Mbakilwa et al. 2007). Levels of resistance also affect results and can vary widely even in close geographical or temporal proximity (Amin, Hughes et al. 2004).

Regarding diagnostics, variation in transmission intensity will result in different prior probabilities of infection, which together with test accuracy determine the test predictive values. Differences in host immunity will also determine the likelihood of infection being the cause of illness and the severity of disease. Where immunity is high there will be lesser benefit in demonstrating parasitaemia. For these reasons results obtained in any single trial are unlikely to be easily generalisable to other settings

The use of numerous centres in RCTs is one method of increasing the generalisability of results. Despite the potential such studies hold, they also pose greater methodological challenges in how the variability between study sites is addressed. In the multi-centre study of Chanda, Masiye et al. (2007), the data from all sites were simply pooled, potentially masking significant site differences in costs and effectiveness.

A number of shortcomings in how results were interpreted were found in the studies presented. The Jonkman study on switching to a strategy of biological confirmation did not include a measure of effectiveness at all (Jonkman 1995). This study therefore, is a cost analysis, not an economic evaluation, as it does not incorporate a measure of outcome other than the final expenditure under each strategy. Consequently, the lower expenditure with the strategy of microscopical confirmation gives no real indication on the implications for patients' health outcome which may be compromised under the strategy. The study could lead policy makers to erroneously conclude that the intervention is justified from an economic perspective.

Secondly, the policy context in which results are to be interpreted needs to be explicitly stated. Both evaluations of diagnostic tests limit their measure of effectiveness to test accuracy rather than final health outcome. Such analyses can be appropriate where it has previously been decided that a positive test is a prerequisite for treatment, as is the case in a number of countries in Southeast Asia and in South Africa. They do not, however, evaluate the benefit of a strategy of parasitological confirmation in terms of final health outcome, and do not allow for a comparison with presumptive treatment or other competing interventions.

Some of the above concerns can be addressed by entering external data and assumptions, so that outcomes of greater interest can be estimated.

2.4.2 Category II: Simple cost-effectiveness analysis using multiple data sources

This category consists of either trial based studies that incorporate additional external data in the analysis, or studies that rely entirely on secondary sources to estimate intervention costs and effectiveness, without the formal use of a decision analytic model. These data are aggregated to estimate incremental cost-effectiveness ratios (ICERs) and other variables of interest (Goodman, Coleman et al. 2000; Pang and Piovesan-Alves 2001; Agnamey, Brasseur et al. 2005; Mulligan, Morel et al. 2005).

The greater flexibility and scope in data sources allows the analyst in these studies to answer broader questions than the efficacy of an intervention in a specific clinical trial. Instead, a wider variety of inputs such as varying disease prevalence, population age structure or the value placed on the outcome of interventions, can all provide results that may be of more interest to decision makers.

The extrapolation to final health outcome, introducing measures such as Disability adjusted Life Years (DALYs) or Quality Adjusted Life Years (QALYs), allows for the incorporation of individual or societal values of the improvements these interventions provide in terms of health related quality of life (discussed in detail in Chapter 3). Extrapolating to final health outcomes also allows for comparisons of a range of different interventions competing for the same resources such as national malaria control programme budgets.

Goodman et al. (2000) in particular provide a comprehensive example of how these cross-intervention comparisons can be achieved. Following the launch of Roll Back Malaria, the review set out to calculate the cost-effectiveness ratios (cost per DALY averted) for a range of malaria related interventions targeting diagnosis, treatment and vector control. The aim of the work was to inform policy makers on how best to prioritise these interventions given the expected increased funding for malaria control. Calculating the costs and

effectiveness for such a broad range of interventions required that the analysis draw on a wide range of published and unpublished material, and where necessary make use of expert opinion. Where such broad analyses are done, these inevitably rely on multiple assumptions where data are lacking, and value judgments which might be contentious (e.g. discount rates for future health gains).

Use of these data sources and assumptions is in stark contrast to trial based evaluations that limit their analyses to parameters for which primary data are available. Although the data and methods used in the former might appear less scientific, they do provide policy makers with the kind of information they need to make appropriate decisions. The challenge in such analyses is to make the uncertainties and assumptions as explicit as possible and handle the uncertainty surrounding parameter estimates adequately. Decision analysis is one of the most frequently used approaches to facilitate this.

2.4.3 Category III: Decision analytic models

The differences between the studies in this category and the previous one are limited to the structure used to converge the different factors accounted for in the analysis. The analyses in this category make use of formal decision analytic models to incorporate all relevant factors to produce an explicit recommendation as to which is the best course of action given numerous choices and uncertain events. Ten studies were identified that took a decision analytic approach, two of which employed additional modelling methods and will be discussed in the next section (Sudre, Breman et al. 1992; Cho Min and Saul 2000; Wilkins, Folb et al. 2002; Muheki, McIntyre et al. 2004; Rolland, Checchi et al. 2006; Zurovac, Larson et al. 2006; Shillcutt, Morel et al. 2008).

One of the primary tools used in decision analytic modelling is the simple decision tree. These have a decision node at the origin (i.e. the policy/clinical options branching out of that point) followed by pathways representing the different possible sequence of events up to defined end points where the costs and benefits can be evaluated for that pathway. The probability of each of these events occurring is entered into the model, based on clinical trial data, secondary sources and expert opinion. Each of the outcomes is assigned

a value such as costs and/or utilities. The different branches are then summarised using both their probabilities and outcomes. The branch with the highest expected payoff will be that recommended for policy adoption.

Figure 2-1 illustrates a simple decision tree aiming to shed light on whether a febrile patient should be treated presumptively, or tested prior to treatment. The square node represents a decision node, the circle ones chance events, and the triangular ones terminal nodes.
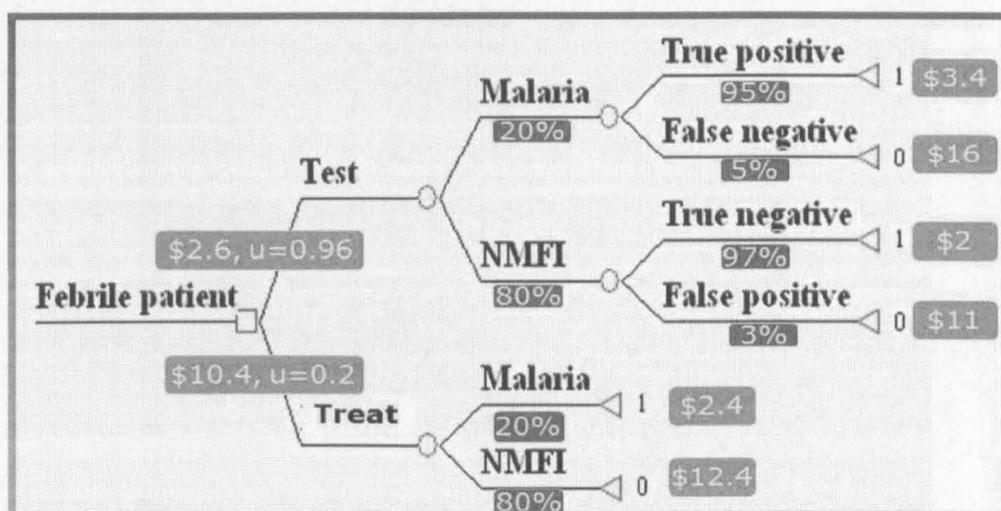
Figure 2-1 - A simple decision tree for the management of malaria suspected patients. The monetary values to the left of each triangular ('terminal') node represents the cost of reaching this point. The utility of each terminal node is fixed as either 0 or 1, dependent on whether the patient has been appropriately treated. These are then 'folded back' to the branches coming out of the decision node to summarize their costs and utilities. In this case the 'test' option, at a lower cost of $2.6 and a higher utility of 0.96, would be preferred. NMFI – non-malarial febrile illness.

This process requires the modeller to explicitly state the probabilities, costs and utilities for each specific uncertain event and outcome. Once these are entered the outcome can be subjected to sensitivity analysis to test any of the values entered. The results are then available to decision makers in a transparent and unambiguous manner (Lindley 1985).

The earliest attempt at conducting a decision analysis on the use of antimalarials was that by Breman, Sudre et al. (1992), comparing CQ to sulfadoxine-pyrimethamine (SP) and to amodiaquine for the treatment of children in SSA, under the circumstances of increasingly prevalent CQ resistant *P.falciparum*. The model included a limited number of variables,

illustrating the simplest of scenarios – a single treatment of febrile children, comparing the outcome for the different drugs in terms of cost per cure and cost per death averted.

A later decision tree evaluation, by Wilkins, Folb et al. (2002), was a more detailed study of the average cost-effectiveness of SP and CQ. Using a range of data sources such as *in vivo* efficacy trials for the drugs in the location of interest (Mpumalanga, South Africa), diagnosis costs and patient travel costs, the authors constructed a decision tree to simulate comprehensive treatment costs for each of the drugs. Treatment failure might lead to the return of the patient to the health care facility (probability assigned after consultation with local health workers) with either severe or uncomplicated malaria, followed by treatment with quinine.

While the advantages of using a decision analytical approach are apparent in these studies, there are a number of issues and possible shortcomings that warrant further discussion.

Model structure – choice of parameter. Trying to include high numbers of parameters results in unwieldy trees that may be difficult to manage. To avoid this, analysts must use their discretion to decide on those parameters that are deemed most significant. This however, opens the model to criticism regarding their subjectivity, in that inclusion/exclusion of some parameters and not others might lead to contradictory results. For instance, choice of costs that follow treatment failure could be seen as arbitrary (e.g. in Wilkins et al. (2002) with the inclusion of administrative and support costs, comprising over half of total recurrent antimalarial costs, but not productivity or home care costs).

How probabilities are obtained. The potential drawbacks of relying entirely on trial data were discussed in an earlier section. Other sources of data for probabilities can be the published literature, or expert opinion. Muheki et al. (2004) demonstrate the use of Delphi surveys to quantify expert opinion for parameter estimation where no empirical data are obtainable. While it could be argued that the use of such methods lacks statistical rigour, they can be employed to address questions that no trial could have answered.

Handling parameter uncertainty. In the Cho-Min and Saul (2000) study, handling of uncertainty surrounding model parameters and structure is limited to a one way sensitivity

analysis for treatment efficacy. In a study based entirely on modelling secondary data, there is considerable scope for introducing a range of likely values, which was not explored in this analysis. In the context of malaria treatment where there are such large elements of uncertainty, these must be allowed to carry through the analysis rather than adhering to simple point estimates.

Wilkins et al. (2002), Rafael et al. (2006) and Shillcutt et al. (2008) all used probability sensitivity analyses (PSA), where probability distributions are assigned to input parameters and Monte Carlo simulations are repeatedly carried out for parameter estimates. In addition to introducing a range of likely values, the use of probability distributions also allows the assignment of distribution densities reflecting the nature of the parameter, for example treatment costs often being highly skewed to the right (Fenwick, Claxton et al. 2000).

In the most comprehensive analysis of its kind, Shillcutt et al. (2008) used a decision tree analysis to estimate the cost-effectiveness of the use of RDT and microscopy in conjunction with ACTs, in comparison to the predominant practice of presumptive treatment. In contrast to almost all other analyses, the authors also incorporate the health outcomes of NMFIs. This introduced in to the analysis what is perhaps as significant a factor as the outcomes for true malaria cases, but also a range of additional uncertainties and assumptions, such as the diagnoses of these illnesses and their health outcomes with and without appropriate treatment. While in their analysis the authors did not have data to estimate these parameters adequately, the uncertainty surrounding them did indicate that this was an important issue requiring further research.

Rafael et al. (2006), although not a proper economic evaluation as it did not account for financial expenditure, is also noteworthy due to the incorporation of the harm associated with the use of antimalarials, placing a necessary constraint on what could otherwise be unrestrained use of drugs.

Linear/one way movement. Decision tree structures imply linearity in the progression of events. This may be far removed from reality. In a number of these studies the decision

tree depicted linear care seeking patterns in the confines of the formal sector. In reality health care seeking behaviour is less manageable as transition between the different states can be far more complex and care seeking patterns often involve the informal sector where the bulk of malaria treatment takes place (Chima et al., 2003).

Choice and interpretation of outcomes. In deciding between the adoption of different interventions, there will seldom be one that dominates the other on all accounts (usually cost and effectiveness). For this reason the choice of measure of outcome can be a decisive factor, as comparing interventions on the basis of either cost minimization, average cost, or incremental costs can all provide different results.

A number of studies in this category used average costs to determine cost-effectiveness. This can be misleading as in some instances an intervention with higher average cost will still be considered cost-effective as decision makers might be willing to accept the incremental cost as a justified investment for the attainment of the incremental benefits provided by the more expensive intervention.

Rolland et al. (2006) and to a greater extent Shillcutt et al. (2008) incorporate in the analysis policy makers' willingness to pay for the benefits the intervention provides. This was incorporated with the use of the net benefit approach, where the results in terms of ICERs are plotted against the policy makers' willingness to pay for the benefits in a probabilistic manner (i.e. the probability the intervention will be cost-effective if the policy maker is willing to pay $X per DALY averted).

2.4.4 Category IV: System dynamic models

Decision trees were developed mostly in the context of clinical options for individual patients at a specific point in time and may be less suitable for long term planning where variations in factors such as seasonality or emergence of drug resistance will alter final outcomes. None of the previous analyses consider, for instance, the dynamic nature of drug efficacy due to growth of parasite resistance to antimalarials, or changes to transmission intensity following changes in control strategies. To incorporate these, models must increase their complexity by adding a temporal dimension. This and other

similarly complex dynamics can be addressed with the use of differential equations and other mathematical operations to simulate the relationships between variables (Mortimer, Smith et al. 2003).

A number of papers were found that attempted a temporal analysis of changing first line drugs, including both the most appropriate choice of drug, and the best point in time for switching to its use (Schapira, Beales et al. 1993; Goodman, Coleman et al. 2001a; Mortimer, Smith et al. 2003; Laxminarayan 2004; Morel, Lauer et al. 2005; Tediosi, Maire et al. 2006; Yeung 2006).

Laxminarayan (2005) produced a mathematical model of malarial transmission, immunity, and drug resistance, which was used to compare the economic consequences of treatment strategies (replacing CQ with ACTs or with SP and later ACTs when resistance to SP develops). The study also explored the economic impact of different levels of coverage with ACT given their high costs and the increasing likelihood of evolving resistance.

The model provided a macro view to identify the best timing for introducing ACTs and the ideal coverage levels. The model, however, was not designed to provide more detailed decision recommendations as it did not account for other factors such as different age structures or varying transmission levels.

Morel et al. (2005) estimated the cost and effectiveness of each of the interventions drawing on a range of secondary data sources, including trial results, expert opinion and the WHO-CHOICE database for regional cost data (Edejer, Baltussen et al. 2003). The effectiveness estimates were based on the WHO PopMod model (Lauer, Rohrich et al. 2003), that uses differential equations to estimate transition of populations between different disease states and evaluating the changes resulting from the introduction of interventions.

This form of analysis does not presume the existence of a 'current practice' as the baseline for an incremental cost-effectiveness analysis, but rather evaluates all interventions, existing or under consideration, for efficiency against a counterfactual situation of doing nothing. The results of the analysis indicated that while the use of ACT on its own is the

single most cost-effective intervention, maximum effectiveness would be achieved by investing in other interventions as well rather than attempting to increase ACT coverage to 100%.

The greater generalisability and scope of these analyses does have a drawback, in that they become increasingly detached from local variation in circumstances; this was particularly true in the case of Morel et al. (2005) with the use of generalised cost-effectiveness analysis. The result is a blurring of the impact of local variation of factors that can be as specific as the time children in the area go to sleep (i.e. protected under insecticide-treated bednets), which could affect parameters of interest significantly. Trying to improve the accuracy of these models to reflect better real life events can require extremely complex structures which are not amenable to further adaptation (Mortimer, Smith et al. 2003).

A further strength of the study by Morel et al. (2005) and the one by Goodman et al. (2001a) (unrelated to the use of system dynamic modelling), is the use of DALYs as the measure of outcome, providing policy makers with a reference point – in this case $25/DALY averted, as a threshold to consider an intervention cost-effective. In addition policy makers' time preferences were also included in these analyses to allow for a higher/lower valuation of more/less immediate costs and benefits.

System dynamic modelling offers significant potential advantages in the ability to account for long term ramifications of the adoption of new health care interventions. They are mostly adequate to assess the broader dynamics of factors such as drug efficacy, emerging resistance, and general coverage levels.

A limitation of these models, which relates to the work presented later in the thesis, is that the higher technical specification of these models can prohibit their adaptation to different circumstances by stakeholders. This trade-off between model complexity and methodological validity is discussed further in Chapter 6.

2.5 *Chapter conclusion*

Despite the overwhelming impact of malaria and NMFIs on population health in SSA, and severely limited resources to manage these, relatively few economic evaluations of malaria treatment and diagnostics were identified, of which only eighteen have been published.

The review opened with those studies grounded entirely in patient level data from clinical trials which estimated of an intervention's cost-effectiveness. At the other extreme were evaluations based entirely on secondary data that made use of a variety of modelling methods to incorporate a range of factors such as emergence of resistance to antimalarials.

Movement along this continuum increased the relevance of results to decision making by broadening the scope beyond immediate costs and effectiveness. However this also introduced additional assumptions and increased the extent of uncertainties. To handle the greater uncertainties involved, a variety of sensitivity analyses were used, as were probability distributions rather than point estimates. A small number of analyses incorporated the likelihood of different results given policy makers' differing willingness to pay for the intervention being evaluated.

General limitations in existing studies. A number of limitations were apparent in the literature. First and most importantly was the need for more responsiveness to different sub-groups, environments and to specific RDTs and treatments, rather than trying to determine whether RDTs or ACTs were or were not cost-effective for SSA as a whole. Secondly, was the need for more comprehensive analyses that included the health outcomes for NMFIs, the broader societal costs of over-treatment with antimalarials, and the costs and effectiveness of interventions as they are likely to manifest in routine practice, rather than trial settings.

Another limitation of many previous evaluations was that they used inappropriate frameworks such as cost-effectiveness analysis using natural measures of outcome (e.g.

cost per malaria case treated or cost per false negative averted). These frameworks were limited in their relevance to decision making as their outcomes were not comparable to other interventions. The ability to incorporate other factors to allow for more comprehensive analyses in these frameworks, such as resistance to antimalarials, was therefore limited.

Lastly, many of the studies were not amenable to use by policy makers because they were setting specific, or conversely overly generalised, not accounting sufficiently for local variation in circumstances. Methods to apply results from one setting to another require further exploration to improve the integration of emerging local data with existing evidence in the models. This can allow policy makers to make better use of existing evidence and models to identify the most efficient interventions and strategies for settings and populations of interest.

3. Theoretical foundations of decision models for malaria diagnostics and treatments

The aim of this chapter is to introduce the theoretical foundations and concepts that underlie decision models, relating to the economic evaluation frameworks they use, how to incorporate data such as costs and effectiveness in the analyses, and how uncertainty is handled. The generalisability of evaluations based on single and multi-centre trials is also discussed. Details of the actual model structures used in subsequent analyses are described in later chapters.

3.1 *Overview of economic evaluation frameworks*

Economic evaluations are commonly used in health systems of the developed world and increasingly so by donors and national governments in developing countries to inform policy makers on the efficiency of new interventions. The basic premise for these evaluations is that finite resources and virtually infinite demand for health care require that different treatments and services are evaluated in terms of their costs and consequences to ensure efficient allocation (Drummond, O'Brien et al. 2005).The concept of opportunity costs is central to this, expressing the notion that any intervention pursued needs to be considered in light of what could have been done with the same resources elsewhere.

Table 3.1 is an adaptation from Morris et al. (2007), showing the different levels of opportunity cost that could be considered when funding is diverted towards a particular health care intervention, in this case the establishment of a home management of malaria (HMM) programme for a district with poor access to health care facilities.

Level	Choice about	Decision to	Opportunity cost
Government, all sectors	How much should be spent on health care in total	Increase spending on health care through higher taxation	Lower net incomes – reduced spending power; higher taxation on producers lowers incentive to innovate and invest in further production; reduced salaries; higher prices as producers 'pass on' cost of taxes to consumers
		Increase spending on health at the expense of other sectors	Benefits forgone from lower investment in education, road safety etc.
Health sector	How much to allocate to remote areas with high malaria endemicity	Increase the share of remote areas from the total health budget	Benefits forgone from other geographical areas with higher population density
District	What share of the district health care budget should be provided to the new HMM programme	Increased spending on HMM through recruitment of more community drug distributors	Benefits forgone from investing in prevention and cure of other illnesses
Programme	Which population and sub-groups are targeted; children of all ages or only under 5 years with lower immunity	Provide HMM services to children up to 10 years	Lower returns on the investment as older children are more likely to be immune. Benefits forgone from treating other illnesses to which they are more susceptible
Patient / carer	How much would carers be willing to spend in cost sharing schemes for HMM	Levy a charge for regular provision of antimalarials	The utility forgone on spending and consumption of other goods and services

Table 3-1: Opportunity costs on the way to funding a new antimalarial distribution programme for areas with poor access to health care (adapted from Morris et al. 2007). HMM – Home Management of Malaria

While in most other fields the market is often seen as the ideal medium to ensure the allocation of resources reflects their true opportunity cost, it is largely accepted that in the sphere of health care, unregulated markets will fail to achieve this. Instead, analytical processes are employed to evaluate interventions' costs and consequences in order to guide efficient allocation (Rice 1997).

Economic evaluations aim to compare alternative interventions, usually proposed new ones to existing practice, by combining estimates of their costs and consequences to assess whether these are a worthwhile use of resources. These estimates can be drawn from trial data, existing literature, or expert opinion, and can then be synthesized using a variety of modelling techniques to extrapolate results to other settings or longer time horizons.

Depending on the context, the adoption of a new intervention might imply forgoing investment in other interventions associated directly with the same aims, or with interventions with other benefits, either within or beyond the health sector. Both costs and consequences can be assessed from a variety of perspectives –those of the provider, the patients, or society as a whole. The context of the necessary decision and the relevant perspective will determine the most appropriate framework for use (McPake, Kumaranayake et al. 2002).

3.1.1 Economic evaluation frameworks

Economic evaluations can be differentiated primarily by the methods used to assess the intervention consequences. While in other fields cost-benefit analyses (CBAs) are the dominant form of evaluation (Robinson 1993), where both costs and benefits are captured using monetary values, it is often perceived as unpalatable to do so in the context of health care. Therefore since the 1960s when CBAs first emerged in health care evaluations (Warner and Hutton 1980), a variety of methods and frameworks have attempted to circumvent the assignment of monetary values to changes in life expectancy and quality.

The simplest alternative is where intervention consequences are measured in natural units, for instance the number of children covered by an insecticide treated bed net

programme (Stevens, Wiseman et al. 2005). In this case the evaluation is classified as a cost-effectiveness analysis (CEA). Results will normally include the average cost per unit of effectiveness, and an incremental cost for an additional unit where one intervention is more costly and more effective than the other (e.g. incremental cost per year of bed-net coverage). It is then up to the decision maker to consider whether the cost of this incremental benefit is an efficient use of resources (Phelps and Mushlin 1991). This process, however, is hampered by the fact that the decision maker will often have to choose between interventions with very different aims and measurements of outcome.

To overcome this limitation, intervention consequences can be measured in generic units, measuring either changes in quantity or quality of life, or a composite of both as in the widely used measures of QALYs and DALYs. An evaluation of ITNs could therefore also measure its benefits in DALYs averted (Coleman, Goodman et al. 1999), allowing a comparison of interventions with entirely different aims. Evaluations that incorporate a measure of quality and quantity of health outcome are classified as cost-utility analyses (CUAs), and have become the dominant form of evaluation in heath care, with QALYs being used primarily in high income countries, and DALYs in low income ones (Drummond et al. 2005).

There is considerable variation in the literature on how to classify the different forms of evaluations. A common alternative to the classification used here is the use of CEA to include all evaluations where outcomes are not measured in monetary terms, with CUA as a sub-category of these reserved for those evaluations that follow very particular methods for estimating utility of health outcomes, confined to measuring the individual utility a patient/society places on the benefit of the intervention of interest. This in practice is only true for evaluations using QALYs as a measure of outcome. In this thesis however CUA is treated as a separate category from CEAs and includes evaluations that use DALYs as a measure of outcome as well. The term *cost-effectiveness*, however, is used in the broader sense, summarising the costs and consequences of an intervention regardless of the framework used.

3.1.2 How results are applied; decision rules

Economic evaluations are carried out with the aim of making informed choices between competing health care interventions. As described above, the different frameworks are distinguishable by the methods used to express the interventions' outcomes. The significance of these differences becomes most apparent when considering how these results are applied in the decision making context. 'Decision rules' are the criteria against which the results of the evaluations are assessed to determine whether an intervention is considered to be cost-effective. A decision rule can state a threshold below which an intervention is considered cost-effective, with or without a measure of uncertainty attached. Such thresholds are often referred to in the literature as the *ceiling ratio,* and in economic analyses' notation are assigned the Greek letter λ.

In CBAs the decision rules are a direct product of the analysis. If an intervention has higher benefits (in monetary values) than costs, the intervention is a worthwhile investment. Where a number of alternative interventions are being proposed these can be evaluated by identifying that with the highest net benefit (benefits minus costs) or that with the higher benefit-cost ratio (BCR).

In contrast when outcomes are expressed in natural units as in CEAs, no clear decision rule is apparent. In this case the analysis can be used to guide choices only between interventions with identical aims, and under the assumption that the budget for these is not competing with that for other interventions (Drummond, O'Brien et al. 2005). As this is rarely the case, decision makers will be left to consider intuitively whether the benefit of an intervention with one aim at a particular cost is more efficient than an intervention with different aims and costs.

CUAs, on the other hand, measure benefits in generic terms that allow direct comparisons between interventions with different aims. In this case decision rules can be applied by deriving thresholds for the cost per QALY gained or loss of DALY averted, below which an intervention is considered cost-effective. There are a number of methods for determining this value, but only one of these, the league table approach, is consistent with the aim of avoiding the assignment of monetary values to health outcome on a normative basis

(Gerard and Mooney, 1993). The other approaches are normative, based on either the human capital or revealed preferences (Sachs 2002; Shillcutt, Walker et al. Unpublished), sharing the theoretical foundations of CBAs.

3.1.3 Determining the ceiling ratio using the league table approach

The league table approach to determining the ceiling ratio requires estimating the ICERs for all existing and proposed interventions. These are placed in a league table and the health budget is allocated starting with the most cost-effective interventions in descending order until the budget is consumed (Drummond, O'Brien et al. 2005). The cost per QALY or DALY of the last affordable intervention is then assumed to be the decision rule[4].

Using this approach, the World Bank and later the WHO developed the commonly cited thresholds for low income countries of $25 and $150 per DALY averted below which interventions are considered 'highly attractive' or 'attractive', respectively (Jamison 1993; WHO 2006c).

The league table approach to relies on a number of assumptions regarding the health budget – that this is a fixed amount, and that funding can be applied interchangeably between interventions, which themselves are divisible. These assumptions, however, might not hold, particularly in the context of low income countries.

First, the fluctuating funding received from donor agencies that often constitute a considerable part of the government's own health care spending, makes it difficult to calculate an enduring ceiling ratio. Secondly, much of this funding may be restricted to particular diseases or interventions, so funding may not be interchangeable between all interventions. Lastly interventions are often not divisible, and will be rolled out in a manner that is highly efficient for some populations and locations, but not for others. In

---

[4] More precisely this would be the cost for the last marginal QALY gained, assuming all interventions are divisible.

addition to these limitations are the immense informational requirements – that ICERs for all proposed interventions be known, as well as the health budgets to fund these.

Where ceiling ratios are used, for them to remain relevant they would have to be regularly readjusted, otherwise the total health budget would continuously increase to accommodate new interventions with ICERs below the decision threshold are introduced (Gerard and Mooney 1993). Furthermore the ICERs that determine an intervention's ranking in a league table will often be highly dynamic, as prices fluctuate and epidemiological realities change. Having static league tables on which decisions are based will therefore lead to allocative inefficiencies. This is particularly true in the context of the developing world where factors such as subsidies or exchange rate fluctuations can have rapid and substantial impact on intervention prices. Static league tables and ceiling ratios cannot easily handle such dynamic environments as they take extensive time and resources to establish and could rapidly become outdated.

3.1.4 Normative approaches to determining the ceiling ratio

The league table approach tries to apply an entirely pragmatic method to determine how much should be spent on an additional unit of health outcome, such as a DALY averted, avoiding *a priori* statements of its monetary worth. The alternative approaches to determining the ceiling ratio differ from the league table approach in that they assess the inherent value placed on, for instance, gaining an additional year of life lived in full health, regardless of the available budget and other constraints.

One method of doing so is through the direct elicitation of either policy makers' or the public's willingness to pay (WTP) for the intervention benefits. Similarly *revealed preferences* can be used, where previous decisions made by policy makers are evaluated to identify the threshold below which an intervention should be adopted (Behrman, Alderman et al. 2004; Drummond, O'Brien et al. 2005). The revealed preference approach led initially to some informal estimates for the ceiling ratios of NICE in the UK (Bryan, Williams et al. 2007). More recently NICE have formally adopted this threshold as a guide to their decisions.

Alternatively a human capital approach can be taken using measures such as GNI/capita to estimate the value of a DALY averted (Sachs 2002; Shillcutt, Walker et al. Unpublished). The benefit of this method is that it is explicitly linked to a country's economic capacity to invest in new interventions. It should be noted that both the human capital approach and the welfare theory on which the WTP approach is based, form the theoretical foundations that underlie CBA. These normative approaches are in contradiction with the desire to avoid the assignment of monetary values to health gains, which was to a great extent the impetus behind the use of CEAs and CUAs in the first place.

3.1.5 Incorporating ceiling ratios in CUAs

Once a ceiling ratio has been established, a variety of methods have been developed to incorporate it with the outcomes of CUAs. The simplest use of the ceiling ratio is by comparing it to the ICER for a QALY gained or DALY averted from the CUA. Where the ICER exceeds the ceiling ratio, the intervention is not considered to be cost-effective. This method, however, carries a number of drawbacks, most importantly that it requires a single explicit value for the ceiling ratio which are seldom readily available (Gafni and Birch 2006). Furthermore, use of a point estimate for the ICER does not capture the variability and uncertainty surrounding the intervention costs and effectiveness. One method that incorporates the uncertainty in the analysis and explores results across a range of ceiling ratios is the use of cost-effectiveness acceptability curves (CEACs). This expresses the probability that an intervention is cost-effective across a range of ceiling ratios, circumventing the need to pin it down to a single value (O'Hagan, Stevens et al. 2000; Fenwick, O'Brien et al. 2004).

A general limitation of the ICER concerns the difficulty in attaching to it measures of uncertainty and the subsequent interpretation of changes to its value and sign (Hoch, Briggs et al. 2002). Figure 3-1 shows a cost-effectiveness plane with the results of a Monte Carlo simulation, where the uncertainty surrounding costs and benefits are obtained simultaneously. While most data points are in the north-east quadrant, indicating a more costly and more effective intervention, there are a smaller number of data points in the other three quadrants.
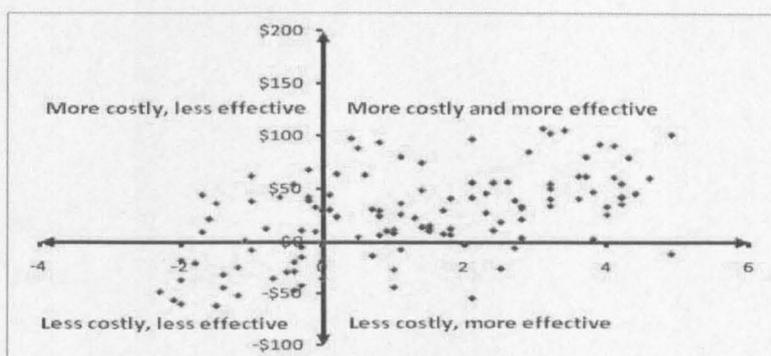
Figure 3-1: A cost-effectiveness plane showing the different possible relationships between the costs and effectiveness of the interventions being compared

As the ICER is a ratio, whether it is positive or negative will depend on the signs of the differences in both cost and effectiveness between the two interventions being compared. Numerically, any single value of an ICER can originate in one of two quadrants. A negative ICER can be a result of the new intervention being cheaper and more effective, or conversely more expensive and less effective. These ICERs therefore cannot be simply averaged out as identical numerical values can represent outcomes with opposing interpretations.

To overcome this, the net benefit framework was developed to provide a linear value where negative and positive results are unequivocal in their interpretation (Stinnett and Mullahy 1998). The net benefit framework requires that the ceiling ratio be explicitly incorporated in the analysis. It uses this value to convert differences in effectiveness into a monetary value that can be fully merged with the difference in cost to provide the net monetary benefit of the interventions. Alternatively the same ceiling ratio can convert the difference in cost into a measure of health gain to estimate the net health benefit. The two methods are entirely equivalent.

While the new measure was initially proposed to overcome the technical limitations in the use of ICERs, it has allowed analysts to employ the powerful toolset of econometrics to economic evaluation. Recent use of econometric methods in economic evaluation has shown they make more efficient use of data and provide greater ability to explore a range of issues including sub-group analyses, interaction, and confounding in a more robust

manner than was previously possible with the use of ICERs (Hoch, Briggs et al. 2002; Grieve, Nixon et al. 2005).

Another consequence of using the incremental net-benefit approach is that this has essentially transformed CUAs into CBAs as the health benefits are now formally expressed in monetary values (Phelps and Mushlin 1991; Stinnett and Mullahy 1998; Drummond, O'Brien et al. 2005).

3.1.6 The WHO-CHOICE framework

In low and middle income countries, CEA frameworks continue to dominate, most notably that of the World Health Organization. In an effort to promote a better evidence-base for its policies and endorsements, the WHO created a framework using standardized CEAs for a wide range of interventions ranked in a single league table (Murray, Evans et al. 2000).

The impetus to creating this framework was partly the fact that despite the theoretical literature's focus on a sectoral approach to economic evaluation, most CEAs in practice compared new interventions to existing ones in an *ad hoc* manner that does not allow for comparisons of interventions to others with different aims, nor for generalisability of results as the baseline interventions used often differed between localities. Furthermore, there was no assurance that the baseline intervention was itself cost-effective, therefore an incremental analysis concluding that the new intervention appeared efficient might only be so due to larger inefficiencies of the baseline intervention. Adopting the new intervention in this case would further worsen existing inefficiencies (Hutubessy, Bendib et al. 2001). Lastly, a positive ICER, even when lower than the decision threshold, implies that adopting a new intervention requires diversion of funding from other existing interventions within or outside the health sector. The opportunity cost of this is not accounted for in incremental analyses (Sendi, Gafni et al. 2002).

As an alternative, WHO recommended the use of generalised cost-effectiveness analysis (GCEA)[5] using average rather than incremental costs per DALY averted as its measure of outcome, based on estimates of how costly and effective the interventions are compared to doing nothing, and ignoring all other local constraints. By building up a database of such standardized cost-effectiveness analyses, the efficiency of interventions can be compared across the entire health sector.

Describing the development of the framework, Murray et al. (2000) discuss the possible uses of such sectoral CEAs. The first is to inform policy makers with fixed budgets of the precise cost-effectiveness ratios of a range of interventions so that these can be prioritized for funding (Murray, Evans et al. 2000). However, they argue that this is overly ambitious as it requires that CEAs account for local constraints, such as pre-existing interventions and infrastructure, and other political constraints if they are to provide applicable results.

Alternatively, they argue that CEAs can take on less ambitious aims and provide more of a general indication of cost-effectiveness, as is the aim of the WHO-CHOICE framework. This is achieved by comparing interventions to a do-nothing approach rather than to existing interventions and relaxing assumptions about other constraints. Results of such analyses can be used to provide a general indication as to whether interventions are 'highly cost-effective, cost-ineffective, or somewhere in between' (Edejer, Baltussen et al. 2003; p.6).

The framework provides a list of guidelines to analysts on how to conduct evaluations to ensure that these maintain standard methods that can ensure comparability and generalisability of results to build up league tables for relevant interventions on a regional basis. Into this league table are introduced a range of ceiling ratios relevant to specified regions, determined by economic and epidemiological similarities, to indicate which interventions are cost-effective. The ceiling ratios were obtained by using a multiple of GDP/capita, based on the work of the Commission on Macroeconomics and Health (Sachs 2002).

---

[5] Note that WHO-CHOICE views CUAs as a sub-category of CEAs

3.1.7 Drawbacks to the CUA/ceiling ratio approach

Determining ceiling ratios using the league table approach is difficult due to the high informational requirements and other limitations described above. The alternative normative approaches have been used in only a few instances with very limited success. In Canada and Australia, for instance, this was done on the basis of reviewing previous decisions and inferring the willingness to pay for a QALY gained. This, however, resulted in static values that did not represent the opportunity costs of marginal resources used (Gafni and Birch 2006). Another consequence of using a fixed ceiling ratio that was observed in these cases was the rapid increase in health expenditure, as had been anticipated in the literature (Birch and Gafni 1992; Gafni and Birch 2006).

More broadly than the methods used to determine the ceiling ratio, there are criticisms of the use of CUAs altogether. Labelle and Hurley, for instance, argued that while CUAs have allowed for comparisons across the health sector at the individual patient level, they are not capable of handling externalities, leading to allocative inefficiencies (Labelle and Hurley 1992).

Another potential impediment to the use of CUAs in low income countries is the lack of an appropriate measure of outcome. Having a utility measure that accurately represents the changes in quality of life as well as life expectancy is a prerequisite for CUAs; conducting these with a flawed measure will compromise allocative efficiency. In the developed world, tools for the measurement of QALYs have been the subject of extensive research for a number of decades. It is widely accepted, for instance, that such measures must be culturally specific, so tools such as the EQ-5D for the measurement of health outcome are tailored to each country where they are employed, and subjected to considerable review (Rasanen, Roine et al. 2006).

The use of DALYs for the purpose of evaluating health gains in CUAs has come under significant criticism as these were developed with an entirely different aim, the assessment of the global burden of disease, and might not be entirely compatible with objectives of economic evaluation (Barker and Green 1996; Gold, Stevenson et al. 2002; Lyttkens 2003; Arnesen and Kapiriri 2004; Mont 2007). For instance the initial methods used to estimate

quality of life in DALYs can be described as crude at best, with both the valuation of health states and factors such as age weighting and discounting being decided on by a small group of experts and then applied across an immense range of populations and circumstances (Mont 2007). Alterations in these value choices change the global burden of disease estimates and subsequent ranking of interventions' cost-effectiveness (Arnesen and Kapiriri 2004).

It can also be argued that the use of this entire framework becomes redundant as the impetus behind it in the first place was the avoidance of placing monetary values on health outcomes. That this is in fact inevitable is demonstrated both in its 'borrowing' of human capital and welfare theory, which form the foundation of CBAs, and in practice with the introduction of methods such as the INB and CEACs to CUAs.

3.1.8 Return of the CBA?

Phelps and Mushlin (1991) argue that the main difference between CBAs and CUAs is in the necessity in CBA to determine in advance a value for the QALY, DALY or year of life lost (YLL), whereas in CUAs the cost per QALY can be reported to decision makers and left for their consideration. However, there is little use in such a measure without a ceiling ratio to indicate whether the outcome is cost-effective. Once a ceiling ratio is introduced with the use of, for instance, the INB method, the frameworks produce almost identical results. CBAs, however, have an advantage in having a more coherent framework as the decision rule is inherent to the analysis and does not require an external value imposed on the results.

A number of additional advantages of CBAs can be identified. Firstly, CBAs are better at handling situations with multiple dimensions of health and non-health benefits. Similarly, the framework allows direct comparisons of interventions with entirely different aims, both within and beyond the health sector (Evans 2004; Mills and Shillcutt 2004). Furthermore CBAs can accommodate consequences for both patients at whom the intervention is aimed and other bodies affected by it. In the context of malaria this would facilitate the incorporation of factors such as the development of resistance to antimalarial

drugs, and non-health outcomes such as improved educational capacities of non-parasitaemic children (Hutubessy, Bendib et al. 2001; Kihara, Carter et al. 2006). Given these advantages it is perhaps not surprising that the proportion of CBAs out of all economic evaluations has been rising in recent years, albeit mostly in the context of high income countries (Hutubessy, Bendib et al. 2001).

The main difficulty with the use of CBAs remains the need to place monetary values on health outcome, with a substantial body of literature discussing the different approaches to this (Birch and Gafni 1992; Pauly 1993; Robinson 1993; McIntosh, Donaldson et al. 1999; McPake, Kumaranayake et al. 2002; Drummond, O'Brien et al. 2005). Despite the different theoretical foundations, empirically it can be concluded that there is a strong correlation between the values obtained from use of the human capital approach based on a multiple of GNI and those obtained from the WTP approach (Shillcutt, Walker et al. Unpublished).

The revealed preferences approach to WTP was described above in the context of ceiling ratios, where previous decisions are used to assess values such as the WTP to avert a loss of a QALY. An alternative to revealed preferences are hypothetical WTP estimates, for which elicitation tools such as various bidding methods are available (Frew 2003). There is, however, evidence to suggest that hypothetical WTP estimates diverge considerably from actual WTP as observed in revealed preferences (Onwujekwe 2001).

The human capital approach also has its weaknesses. The use of measures related to income and productivity raises concern around the exacerbation of existing inequalities in health and wealth, as it implies that an intervention with the same health outcome can be perceived as efficient in a country with high income, but not in one of lower income (Shillcutt, Walker et al. Unpublished, Mills and Shillcutt 2004). This contrasts with the use of universal values such as the $150/DALY threshold that has been applied across the low income countries.

While this argument has a strong intuitive appeal, at least where decisions are made in a highly centralised environment, it does not hold as much strength in a decentralised one, where a universal value for health outcomes will result in the misallocation of scarce

resources. As discussed in Chapter 1, there has indeed been an increasing trend toward decentralisation in malaria related decision making. This is evident in the establishment of national malaria control programmes across SSA, and country-based applications to The Global Fund and other large donors. Use of a value associated with productivity ensures that countries as vastly different economically as South Africa and Burundi have appropriate indicators for how cost-effective and affordable an intervention is for their own purposes. A universal value on the other hand would be inappropriately high for some countries, resulting in unaffordable interventions being pursued, or too low in higher income countries, leading to the abandonment of interventions that could be adopted.

3.1.9 Relevance of the economic evaluation frameworks to malaria diagnostics and treatments

There are a number of characteristics of the malaria context that require consideration when choosing the relevant framework for economic evaluation. The first factor to be considered is the source of funding for these interventions. The rise of malaria on the political agendas, both locally and internationally, have assured a considerable increase in funding dedicated to malaria specifically, and often to particular aims such as provision of antimalarials or vector control. This is true for both foreign donor aid, and also for local health financing as most governments in SSA have separate malaria control programmes with their own assigned budgets.

The implication of this is that the main concern becomes technical, rather than allocative efficiency, as the yardstick is how the costs and benefits of these particular interventions compare amongst themselves and not with others across the health sector (Hutubessy, Bendib et al. 2001). This diminishes the advantage of CUAs as the comparison to other interventions is not as relevant, leaving CEAs and CBAs to be considered. CEAs have no inherent advantage over CBAs, other than avoiding placing monetary values on health outcomes.

A further advantage of CBAs over the other frameworks is the ability to address a wide range of factors beyond immediate costs and benefits of interventions that need to be

accounted for in their evaluation. CEAs based on direct health expenditure and DALYs (or other health measures) alone might fail to capture broader implications of the intervention.

Despite the advantages of CBA, as shown in Chapter 2 this framework is rarely used for the evaluation of malaria diagnosis and treatment, with most analysts opting for either CEA or CUA frameworks. The WHO-CHOICE framework is a particularly attractive option as it provides a readily available set of tools for analysts to evaluate the cost-effectiveness of new interventions. There are, however, a number of limitations in the use of the WHO-CHOICE framework in for the evaluation of malaria diagnostics and treatments.

First, the framework provides a valuable tool for assessing whether an intervention is likely to be cost-effective in relation to competing ones in a very generalised manner. This can prove useful only at early stages of consideration of new interventions applied on a large scale (Edejer, Baltussen et al. 2003), for instance, the overall use of diagnostic tests as opposed to presumptive treatment. The framework, however, is not designed to provide specific recommendations such as the best choice of antimalarial or diagnostic test at specific sites or for particular target populations.

Secondly, given the tied-funding often allocated for malaria related interventions, the framework is ill-suited to assess whether these are cost-effective, as it assesses them in relation to other interventions across the health sector. Consequently malaria interventions that would not be considered cost-effective by the WHO-CHOICE and other frameworks may still be pursued, given the tied funding and political imperatives to reduce the burden of malaria in SSA and other areas affected by the disease. Lastly, the framework does not facilitate the incorporation of non-health benefits, such as improved productivity and educational achievement amongst non-parasitaemic individuals.

So whereas it can be argued that the WHO-CHOICE framework can play a valuable role at the early stages of consideration of new interventions, it is not appropriate for use at the deployment stage, where decision makers must choose between competing interventions, or whether to target these for particular subgroups. All methodologies of course carry

their own limitations; analysts must ensure they employ those that are most relevant to the decision making needs.

This chapter has so far reviewed the economic evaluation frameworks relevant to malaria diagnostic and treatments. The remainder of this chapter provides a brief summary of other major theoretical issues relevant to decision modelling, as a background to subsequent analyses. These consist of the handling of uncertainty surrounding decision models; the use of Bayesian inference to incorporate external data and prior beliefs into the analysis; the generalisability of model output to other settings; and the interpretation of variation in costs and effectiveness of interventions between sites in multicentre trials.

3.2 Different levels of uncertainty in modelling and how these are handled

Spiegelhalter and Best (2003) described a classification of the uncertainties surrounding cost-effectiveness models by distinguishing between different types of model inputs and parameters. In their taxonomy, shared by Briggs (2000), first order variability refers to *chance variability* in a homogenous population. This is inevitable and of little interest as the *expected* outcomes will be an expression of an average value for these.

Heterogeneity concerns between individual variability resulting from either known or unknown patient characteristics. Once these are identified they need not be treated as random variables, but rather should be varied systematically to determine changes in outcomes such as the intervention's cost-effectiveness in relation to different sub-groups.

Parameter uncertainty can relate to either 'state of the world' parameters for which a value could in theory be determined from the data but this is either entirely or partially unavailable, or to assumptions which express judgments on parameters that cannot be estimated from the data, often due to their highly subjective nature (e.g. time preferences for discount rates; values placed on different health states).

State of the world parameters are best handled either using a series of one way sensitivity analyses, or by using probability distributions to express the range, shape, and best estimate for their values (Briggs 2000). Probabilistic sensitivity analysis (PSA) uses Monte Carlo simulations to sample these distributions simultaneously and obtain a measure of outcome that contains all their variation. The use of such distributions can be a better representation of what is known and not known of the parameter as it expresses the possible variation as well as its mean or alternative measures of central tendency. Assumptions on the other hand should always be dealt with deterministically to compare results when these are varied.

Ignorance concerns the lack of knowledge on parameter interaction. This would lead to inappropriate model structuring, misrepresenting the dynamic of the problem. The best way of handling this is to work with a number of model structures in order to compare results and explore the variation between them (Briggs 2000).

Model and methodological uncertainties. It is standard procedure in economic evaluation to explore parameter uncertainty, with the use of PSA gradually becoming the norm instead of one and two way analyses, in high income countries (Claxton, Sculpher et al. 2005). There is, however, a higher level of uncertainty, relating to the evaluation framework and structures used in designing the model that is less frequently questioned, even though this can have a greater impact on results (Brisson and Edmunds 2006).

Methodological choices such as different economic frameworks or whether a discount rate is applied and its value where used, can lead to conflicting results (Cohn 1973). Similarly choice of comparator will determine the incremental cost-effectiveness of a new intervention. Lastly the model structure (decision tree, Markov model etc.) will also determine the estimated costs and benefits of the interventions. While it will often be impractical to test results for all frameworks and model structures, analysts should be aware of how these influence results and provide justifications for the choice of each.

3.2.1 The use of Bayesian inference to incorporate external data and prior beliefs into the analysis

The difference between Bayesian and classical inference is often summed up in that classical inference will state the likelihood of observing particular data (e.g. treatment A was twice as efficacious as treatment B) given a certain hypothesis (e.g. there is no difference between the two treatments). Bayesian inference on the other hand, will state the probability of a certain hypothesis being true (e.g. treatment two is twice as effective) given particular data, such as a trial result showing one treatment being superior to its comparator (Berry and Stangl 1996). The latter assertion is far more applicable to decision making needs as it is a direct statement of the probability associated with the hypothesis of interest. Classical methods, on the other hand, use more convoluted statements of probability that are not amenable for use in decision models (Spiegelhalter and Best 2003).

The 'price' for this extra clarity is that analysts must explicitly state the prior estimates for hypotheses of interest, before the introduction of new data. This belief is then updated using data emerging from trials to obtain a posterior distribution of the parameter of interest. These prior estimates, in the eyes of many classical statisticians, are the main drawback of the Bayesian approach, as they are not necessarily based on trial results that showed statistical significance, but instead might draw on external data such as the broader literature, other studies, or expert opinion (Spiegelhalter, Myles et al. 2000).

Bayesian methods can be particularly relevant for a temporally dynamic situation where parameter estimates will need continuous revision (emerging resistance; reduced transmission) and across locations with different epidemiological or demographical circumstances, where new data can be used to update prior estimates, rather than interpreting them in a void (Sculpher, Pang et al. 2004).

One of the most natural applications for a Bayesian approach is in the evaluation of diagnostic tests, where the sensitivity and specificity of the tests can be obtained in clinical trials, but the predictive power of the tests can only be determined by combining these with the assessment of the underlying probability of a patient having the disease of interest. A more controversial use of a Bayesian approach would be the initial

77

determination of the test accuracies by combining prior beliefs on their sensitivity and specificity, with new data on these emerging from trial results.

3.2.2 Generalisability of economic evaluation results

Economic evaluations are often based on a single trial from a particular setting, and then seek to generalise results to other settings. The data input into these evaluations can consist of intervention costs, effectiveness, target population characteristics, similar data for comparator interventions, and the value placed on outcomes. The values of these input parameters, however, are rarely completely fixed and known. Some of the uncertainty stems from limited data, as described above. Some variation, however, will be due to heterogeneity, i.e. genuine differences in intervention costs and effectiveness, and other characteristics of the location and target population (Briggs 2000; Briggs, Sculpher et al. 2006). For instance, in the context of malaria, transmission intensity can vary widely within small areas (Ye, Kyobutungi et al. 2007), resulting in significant implications for population susceptibility to infection and the predictive values of diagnostic tests.

Input parameters can also vary considerably over relatively short periods of time, such as the effectiveness and costs of antimalarials. Use of evaluations founded on outdated values can result in inappropriate policies being pursued. In Tanzania, for example, over a period of less than 6 years, the drug policy for treating uncomplicated malaria was changed from one relatively cheap but failing drug to another, as some studies failed to use the most recent data for the new drug's efficacy, while others showed the newer drug was already failing. Subsequently policy was changed again to a more effective combination therapy (Kindermans 2004; Mubyazi and Gonzalez-Block 2005).

In trying to ensure the relevance of any one particular evaluation to other settings, a growing body of literature is dedicated to identifying methods of improving the generalisability of economic evaluations (O'Hagan, Stevens et al. 2000; Sculpher, Pang et al. 2004; Drummond, O'Brien et al. 2005; Manca, Rice et al. 2005). One approach is to broaden the range of input estimates to accommodate location heterogeneity. Another approach has been to seek only general indications of an intervention's efficiency rather

than clear decision rules, leaving policy makers to account for relative efficiency amongst a broader range of considerations. While such an approach might be useful in the early stages of policy deliberations, it inevitably provides less precise information when considering a specific intervention in a particular setting. An alternative approach, therefore, is to develop models that facilitate the incorporation of data obtained in a specific setting providing tailored results relevant specific to different settings (Briggs 2000; van Gool, Gallego et al. 2007, English and Scott 2008).

An area of particular relevance to the question of generalisability is how to handle data from multi-centre trials. Part of the incentive to conduct such studies, in addition to the greater number of participants, is that multi centre studies can claim greater generalisability, as their results are not limited to a particular area. While the availability of data from a range of sites does offer a better representation of how the interventions compare in varying circumstances, this also poses methodological challenges in how the variability in outcome is interpreted. Pooling the data is often done, without consideration of the variability between sites (Chanda, Masiye et al. 2007). On the other hand, completely stratifying results fails to take full advantage of the data to produce a single cost-effectiveness result. An alternative approach is the use of multilevel models. As described in Chapter 7, such models draw on individual observations from all sites, while recognizing their hierarchical nature (Manca, Rice et al. 2005).

3.3 Chapter conclusion

Economic evaluations and decision models offer a range of tools that can potentially indicate when and where new interventions are an efficient use of scarce resources. This chapter has presented the main theoretical foundations of the available economic evaluation frameworks and the approaches to handling data and the uncertainty surrounding these in decision models.

It has been argued that the use of CUAs and ceiling ratios requires considerable investment in the development of an adequate measure of utility, in obtaining estimates of all interventions' ICERs, and in obtaining a ceiling ratio to be used as the decision rule. While a

number of developed countries are in the process of establishing their decision rules using methods akin to the league table approach, and others have tried and failed, it is unrealistic to do so in the highly resource constrained and unstable environment in much of the developing world.

The generalised framework developed by the WHO does allow for some simplification of this process, but can be argued to be relevant only at early stages of evaluation, not for more practical, country-level choices concerning specific interventions and target populations. Other factors such as high fluctuations in funding, much of it targeted at specific diseases and interventions, also imply that the CUA framework and league table approach in particular are problematic in the malaria context.

The gradual shifting of CUAs toward CBAs in both theoretical foundations and practical methods has been shown, and the inevitability of assigning monetary values to health outcomes has also been argued. CBAs, it has been argued, can offer more coherent and comprehensive analyses, in particular where intervention consequences might affect a wider population than just the patients at which interventions are aimed, and where intervention benefits have non-health dimensions.

The dynamic and variable circumstances prevalent across malarious regions pose a number of challenges to how data from any particular setting is applicable at different places and times. Such dynamic environments pose considerable challenges to the traditional approach of economic evaluation and dissemination of results. Especially in highly resource constrained environments, it is imperative that decisions be based on the most up-to-date and locally relevant data, if sustainable policies are to be pursued. The use of localisable decision support tools is one possible method to ensure that decision recommendations are tailored to different settings. For multi-centre trial based evaluations, the use of multilevel modelling can facilitate the merging of data from different sites into single estimates for costs and effectiveness, while still accounting for between site heterogeneity.

4. Thesis Aims, Objectives and General Methods

Previous chapters have reviewed the various frameworks and models used in economic evaluation of diagnostics and treatments for malaria, identifying their limitations in providing relevant decision recommendations. This chapter provides an overview of the aims and objectives of the thesis studies, and briefly outlines the methods and models developed in this thesis; these are presented in greater detail in subsequent chapters.

4.1 Thesis aims

The thesis seeks to improve the quality of economic evaluation of malaria diagnostics and treatments by introducing recent methodological innovations drawn from the developed world literature. There is a pressing need for these methods, as Chapter 2 showed that economic evaluations often claim generalisability beyond what is supported by the data, and are insufficiently responsive to local circumstances. Similarly, in rapidly evolving environments, evaluations can quickly lose their relevance, unless the most up to date information is used to modify their results. Economic evaluations of malaria diagnostics and treatments have also often excluded factors that are highly significant to the decision making process.

The thesis aims to develop decision models that incorporate a broader range of factors than immediate costs and effectiveness into the analysis, factors which should be accounted for when considering the adoption of malaria diagnostics and treatments. The thesis also aims to ensure that decision models account for local variations in factors such as malaria epidemiology and antimalarial effectiveness.

4.2 Main objectives

The studies in the thesis develop new approaches to modelling the cost-effectiveness of malaria diagnostics and treatment with the following objectives:

- To explore parameters that are of significance to decision making on management of febrile patients that have previously been ignored, through an evaluation of RDTs that expands to include the degree of adherence to test results and the potential harm associated with provision of ACTs

- To ensure that decision models are relevant to local settings, by constructing decision support tools for the evaluation of RDTs and home management of malaria programmes

- To ensure that multi-centre RCT based economic evaluations are sensitive to local variation in trial results, through the use of multilevel modelling applied to a trial comparing the use of quinine and artesunate for the treatment of severe malaria

- To improve the portrayal of disease progression and treatment in decision models, through the use of a Markov model for evaluating home management of malaria programmes.

The general methods used to achieve these objectives can be summarised under three headings – the economic evaluation framework used, the factors included in the models, and the model structures used to assess the impact of the interventions being assessed.

4.3 The CBA framework

In Chapter 3 the different frameworks used in economic evaluations were discussed, and a number of arguments were presented in favour of using cost-benefit analysis in general and in the context of malaria in particular. Firstly it was argued that it is difficult to avoid valuation of health benefits in the decision making context. Use of CEAs/CUAs does not absolve policy makers from the need to decide on a threshold above which an intervention is not cost-effective, but merely results in them doing so in an implicit manner.

CBAs on the other hand allow for the direct incorporation of this value in the analysis. There are a number of advantages to this. First the use of a single unit to measure both costs and benefits allows for a more coherent analysis. Secondly, the use of CBAs facilitates the comparison between interventions with different aims, within and beyond the health sector. Lastly the use of monetary values allows for the potential incorporation of non-health benefits. This is of particular importance with regards to malaria, as the impact of the disease goes far beyond the immediate health outcomes (McIntosh 1999).

4.3.1 Choice of method to assign monetary values to health benefits

While CBAs offer a more comprehensive and coherent framework for economic evaluation, the need remains to determine the value placed on health benefits, a similar task to that of determining the ceiling ratio in CUAs.

In the analyses in this thesis, two monetary values are used for a year of life lost, one being the $150 threshold used by the WHO and the World Bank in the context of DALYs, and the second using GDP per capita as the decision threshold. The World Bank $150/DALY threshold is derived from the World Bank's attempt at defining a set of cost-effective interventions relevant to low and middle income countries, termed the Minimimum Care Package (Jamison 1993; WHO 1996). While the use of the threshold is appealing due to the simplicity of having a single measure for low income countries, there are a number of limitations in the way it was derived. These include the limited number of interventions considered in the MCP and the variation in comparators used when assessing them (Shillcutt, Walker et al. In press). Having a fixed threshold also ignores the variability in affordability between the different countries.

The second threshold used, a multiple of GDP per capita, has stronger theoretical foundations, being based on the human capital approach, following the recommendations of the Commission on Macroeconomics and Health (Sachs 2002). While in theory this reflects individual country capacity to afford different interventions, in practice lower income countries tend to devote a smaller proportion of their GDP to health (WHO 2007c). The higher threshold implied by the use of GDP/capita could result in unaffordable interventions being pursued.

The WTP approach (using revealed preferences) was not used because it requires a consistent decision making process from which an estimate of the value placed on health outcome can be drawn. Given the haphazard nature of health care funding in much of the developing world, this is likely to result in inconsistencies in the WTP for health outcomes (however the lack of empirical data on WTP does not allow for a demonstration of this).

4.3.2 Total costs as a summary measure

The main ways in which the costs and benefits of an intervention can be summarised in a CBA are net benefits (benefits minus costs) and cost/benefit or benefit/cost ratios. Most of the subsequent analyses in this thesis use a value representing the total costs that arise following the adoption of an intervention as the summary measure, as performed in a small number of other economic analyses of malaria treatments (Laxminarayan 2004; Yeung, Pongtavornpinyo et al. 2004). This value incorporates both the direct expenditure on the intervention and other related care, and the costs associated with the intervention consequences. For example, patients with malaria wrongly diagnosed or inappropriately treated have a higher probability of developing severe illness, and these have a case fatality rate assigned reflecting the patient's susceptibility, given their age and transmission intensity. If the patient dies, each year of life lost (the difference between their age and relevant life expectancy) is accounted for by the monetary value attached. This value is added on to the intervention costs providing a value that represents the total cost of the intervention.

The differences between the summary measures for CBAs are not substantial; the reason for using total costs was that net benefit is useful when an intervention is compared to a baseline of doing nothing, in which case the evaluation determines whether the intervention benefits outweigh its costs. In the malaria context, however, presumptive treatment is already in place as standard practice across most of Africa and is the baseline against which the introduction of RDTs in considered. For this reason it is more practical to assign monetary values to all adverse consequences associated with each strategy, add these to their direct costs, and determine which is the most efficient based on that with the lowest value.

4.4 Broadening the range of factors in decision models

As shown in Chapter 2, one of the main limitations of almost all previous evaluations of malaria diagnostics and treatments is that they accounted only for a limited number of factors, usually the immediate costs and intermediate outcomes, ignoring a broader range of costs and consequences that decision makers might want to consider. The main reason for doing so is probably that broadening the analysis will often require drawing on a variety of sources, sometimes subjective ones such as expert opinion, rather than drawing data from a single trial quantifying immediate costs and benefits. Thus analysts might be reluctant to extend their models, fearing that this might render their work 'unscientific' as estimates used will not always live up to various standards of statistical rigour.

Consequently, analyses are often structured around information that is readily available, rather than that which is relevant to the decision makers' needs and considerations. Furthermore, excluding factors that are difficult to estimate does not render them any less relevant to the analysis, but rather places an implicit value on them, and often an extreme and unlikely one. This thesis attempts to draw in a number of factors that have been ignored in most previous analyses, despite their high relevance to decisions concerning malaria diagnostics and treatments.

4.4.1 Adherence to diagnostic test results

Economic evaluations must focus on intervention effectiveness rather than efficacy if they are to provide practical decision recommendations. One of the main potential pitfalls to assessing an intervention's effectiveness pertains to the degree to which it is appropriately utilized in practice once it has been put in place (Drummond, O'Brien et al. 2005). In some instances economic evaluations have used a fixed value that is assumed to reflect the degree of utilization of an intervention, such as 80% capacity utilization in the WHO-CHOICE framework (Edejer, Baltussen et al. 2003).

Evaluations of the cost-effectiveness of rapid diagnostic tests have never accounted for the degree to which clinicians act in a manner consistent with the test result. Such a parameter

could be difficult to quantify, and is likely to vary widely in different contexts, which might explain why analysts chose to ignore it. The exclusion of this parameter, however, is implicitly stating that clinicians' actions are always consistent with test results, a very strong assumption, and as most evidence suggests, an incorrect one (Amexo, Tolhurst et al. 2004; Ndyomugyenyi, Magnussen et al. 2007; Reyburn, Mbakilwa et al. 2007). An evaluation that concludes that RDTs are cost-effective without accounting for the degree to which they are not adhered to will misinform policy makers on the desirability of their deployment. If policy makers are made aware of the impact of varying degrees of adherence on the efficiency of RDTs they could also consider appropriate action, such as the implementation of training programmes for clinicians.

In Chapter 5 an economic evaluation is presented, comparing RDTs, microscopy and presumptive treatment for the management of febrile patients. The model evaluating these strategies demonstrates how results vary in accordance with different levels of adherence to negative test results, such as those obtained in the trial from which the model drew much of its data.

4.4.2 Incorporating adverse treatment outcomes

Another factor, excluded from all analyses of diagnostic tests so far, is the long term costs associated with indiscriminate provision of antimalarials, particularly those associated with antimalarial resistance. The main incentive for considering the use of RDTs is precisely the drive to reduce unnecessary provision of ACTs. The absence of the adverse consequences of using ACTs from evaluations of diagnostic and treatment strategies therefore clearly leads to partial results with limited decision relevance, as the main advantage of their use is excluded from the analysis. Introducing a value for the long term costs of dispensing ACTs can ensure that evaluations account for factors such as the possible development of resistance to ACTs, with both adverse health outcomes and the costs of introducing replacement drugs. Similarly, there are possible long term adverse outcomes due to drug toxicity resulting from consumption of artemisinin derivatives by pregnant women, infants and young children (Price 1999).

The model presented in Chapter 5 allows for different estimates of the adverse outcomes associated with the use of antimalarials to be included in the economic evaluation of diagnostic tests.

4.4.3 Capturing patient costs

While the literature on methods often recommends that economic evaluations take a societal perspective in their analyses (Gold, Siegel et al. 1996; Drummond, O'Brien et al. 2005), in practice all but three of the malaria related evaluations used that of the provider (the exceptions are Laxminarayan 2004, Tediosi et al. 2005, and Yeung 2006). The costs included in all other analyses are only those relating to direct health expenditure by the provider, and not those pertaining to the patient. The reasons for the exclusion of patient and hence societal costs might be the higher informational requirements and the difficulties in estimating factors such as time spent on tending the illness and valuing its worth.

Despite these difficulties there is a strong argument in favour of incorporating what estimates are available in models, particularly when the competing interventions can have different impacts on time demands and financial expenditure for patients and their carers. The difference between using an RDT that can be performed almost instantaneously and sending a patient to a lab for a blood slide might mean little in terms of provider expenditure, but often requires several hours' wait or even a return visit for the patient and their carers.

In two of the models presented, on the evaluation of RDTs (Chapter 6) and of the HMM programme (Chapter 8), users can view results using the different perspectives and gain better insight into the impact the interventions have on patients, providers or society at large.

4.4.4 Health outcomes for NMFIs

In many evaluations the distinction between true malaria cases and NMFIs has not been made, with a tendency to overestimate the prevalence of malaria at the expense of other causes of illness. In such instances the effectiveness of antimalarial interventions are overrated, and the health outcomes for non-malarial patients ignored. While in clinical practice a clinician's presumption that a fever could be malaria might be justified, at higher levels of decision-making and analysis it is imperative that the probability that patients have other causes of illness, and the consequences of not treating these, be accounted for. The challenge this poses however, is that data on the precise breakdown of NMFIs is severely limited, therefore introducing this into analyses increases the uncertainty surrounding the true benefit of the interventions.

In the analyses carried out in this thesis the probability that a patient's true cause of illness is malaria forms the starting point for the models. This allows for greater understanding of how justifiable the interventions are in relation to different transmission areas and levels of host immunity. Different probabilities are also assigned to the prevalence of bacterial illness and the probability that these illnesses become severe; these probabilities are also responsive to patient age, influencing the impact of malaria overdiagnosis on health outcomes in different age groups.

One factor which was not included in the models in this thesis, although is indeed recognized as having potential impact on the costs and benefits of the interventions being evaluated, is the impact of the interventions on transmission intensity. Since the Garki Project, which demonstrated that even the best coordinated attempts at reducing transmission in highly endemic areas in SSA were unsustainable (Molineaux and Gramiccia 1980), there has been a consensus that in these areas transmission is unlikely to be affected by the rollout of malaria control interventions. As the geographical focus of this thesis is mostly SSA, it was decided that this factor would be excluded. This consensus is, however, being challenged by the increasing reports of reduced transmission in many areas in SSA (Greenwood et al. 2008; Ceesay, Casals-Pascual, et al. 2008). This has been hypothesised to be related, for instance, to the widespread presumptive treatment of all

febrile illness as malaria, essentially delivering an intermittent preventive treatment strategy (Gosling et al. 2008).

4.5 Approach to modelling

The third area in which this thesis seeks to improve on existing evaluations is in the model structures used, addressing the following issues:

- The representation of the patient and disease progression paths

- The adaptability of the models to different locations and target populations

- The incorporation of prior beliefs and external data in trial based evaluations

- The generalisability of multi-centre trials and studies.

4.5.1 Patient progression paths – interaction with profiles

The first structural change to standard decision tree structures was the introduction of dependence between parameters, so that factors such as a patient's age or transmission intensity can influence the probabilities for the development of severe illness and case fatality rates. This is achieved by entering estimates, for instance for the development of severe illness for untreated malaria by age and transmission intensity, in a data table, which the model then refers to once a particular age and setting is entered in the model.

As discussed in Chapter 2, one of the limitations of decision trees is their ability to handle only relatively simple dynamics such as single events with linear movement. In the context of managing malaria suspected patients, disease progression may not follow a linear pattern, predominantly due to recrudescence or re-infection, and the development of immunity. Portraying these dynamics in a model requires alternative structures, such as a Markov model.

Secondly, the assumption is simplistic that patients utilize only formal health facilities when experiencing symptoms associated with non-severe malaria, and will re-attend when

facing treatment failure. In fact, care seeking patterns often include home treatment for minor symptoms and alternative treatment for more severe illness, either simultaneously or instead of attending a formal health facility (Biritwum, Welbeck et al. 2000).

A possible benefit of using antimalarials is their provision of a prophylactic effect which can reduce the probability of future episodes once the drug has been administered to treat a febrile illness (White 2005; Gosling et al. 2008). As the treatment of one febrile episode is likely to influence subsequent episodes, a decision tree will fail to capture the full consequences of using different antimalarials.

Markov models have not previously been used in the context of economic evaluations of malaria treatments/diagnostics. As figure 4-1 demonstrates, using such a structure allows patients to shift between the different states (the arrows indicating the movements allowed) according to a transition matrix of the probabilities for each possible movement. Such structures can be used either in the place of particular nodes in a decision tree, or in place of entire branches.
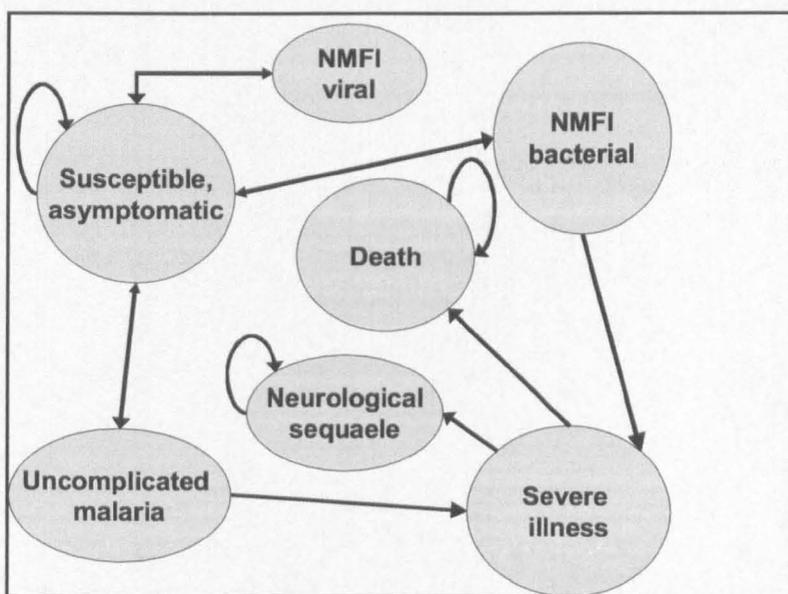
Figure 4-1 - Illustration of a Markov model for malaria suspected patients

In this thesis a Markov model is used to assess the value of home management of malaria with ACTs, as opposed to current practice of seeking care at either private or public health facilities. Given the repetitive nature of febrile episodes and treatment seeking, use of a Markov model is a better representation of this reality.

4.5.2 Localising decision models instead of generalising results

The development of localisable decision support tools (DSTs) could allow for the incorporation of parameter estimates obtained in the target setting to derive location specific results (van Gool, Gallego et al. 2007; Briggs 2000; English and Scott 2008).

The structure of these models is similar to those used in most standard evaluations, for instance a decision tree or a Markov model to assess the costs and consequences of different policy options. The novelty of the DST approach is in making 'user friendly' models available, to permit adaptation of the analysis to other settings (Briggs 2000). Use of such models can also serve as a tool to better engage policy makers and other stakeholders in the analysis. This process in itself can increase the validity of some parameter estimates and might also increase policy makers' confidence in subsequent decision recommendations.

The deterministic model in Chapter 6 and the model evaluating HMM programmes in Chapter 8 were both designed as DSTs and are available for use by stakeholders[6].

4.5.3 Use of a fully Bayesian approach for the evaluation of malaria diagnostic tests

In Chapter 6, a Bayesian decision model using WinBUGS software is compared to a deterministic model for the evaluation of RDTs. The Bayesian model allows for the incorporation of prior beliefs on the accuracy of the different tests prior to the introduction of new data. This can potentially enhance policy makers' confidence in the

---

[6] Information on the models is disseminated through publications in peer reviewed journals and presentations in relevant conferences; the models are available for download with further instructions on their use.

results, by allowing their own prior opinions to be merged in the analysis along with new data as this becomes available.

4.5.4 The use of multilevel models to assess the generalisability of multi-centre studies

The difficulties in interpreting multi-centre trial results were briefly described in Chapter 3. Recently, the use of multilevel modelling has been advocated for use in economic evaluations to address these challenges (Pinto, Willan and O'brien 2005; Grieve, Nixon et al. 2005; Rice and Jones 1997). While this has never been done in the context of malaria, there is considerable need for the use of these models where high variability in costs and health outcomes can be expected, as has been argued to be the case in this context.

In Chapter 7, a multi centre trial comparing treatment of severe malaria is evaluated by comparing the use of multilevel models to standard methods.

The frameworks and methods used in the analyses are summarised in Table 4-1.

	APPROPRIATE FRAMEWORK	ADDITIONAL FACTORS	APPROACH TO MODELLING
Chapter 5: **Broadening the range of factors in decision models**	▪ Use of CBA ▪ Total cost as measure of outcome	▪ Adherence to test results ▪ Harm of treatment ▪ NMFI outcomes	▪ Decision tree responsive to patient age, transmission setting
Chapter 6: **The localisation of economic evaluations using DSTs**	▪ Use of CBA ▪ Total cost as measure of outcome	▪ Adherence to test results ▪ Harm of treatment ▪ NMFI outcomes	▪ Decision support tool ▪ Decision tree ▪ Use of MCMC
Chapter 7: **Multilevel modelling for severe malaria treatments**	▪ Cost analysis		▪ Multilevel regression modelling
Chapter 8: **Markov model for the evaluation of ACTs in HMM programmes**	▪ CBA ▪ Benefit-cost ratio as measure of outcome	▪ Degrees of access to health care ▪ Harm of treatment ▪ NMFI outcomes	▪ Markov model ▪ Decision support tool

Table 4-1: Summary of frameworks and methods used in the analyses. NMFI – Non malarial febrile illness

4.6 Data collection

In order to populate the models developed in the thesis with real data, collaborations with recent and ongoing trials were created, where in return for assisting with the economic analyses of the trials, full access to the data was granted for the purpose of the thesis. Three of these trials took place in SSA and one in Southeast Asia.

As is often the case with clinical trials, the economic component was not a central one, therefore trial designs were not always ideal for capturing costs or outcomes relevant to routine circumstances. To compensate for this, three field trips to Tanzania and Uganda were carried out for prospective and retrospective data collection, from a variety of sources in addition to the actual trial data. These included, for instance, interviews with MoH officials, hospital and clinic costing, and the assessment of time and resources spent by community drug distributors on treating malaria episodes. In the one instance where the trial was still at an early stage when the collaboration was established, an active part was taken in its planning and in data collection and entering. The trial in Southeast Asia was completed in 2005 and the analysis developed here was based entirely on secondary data.

4.7 Chapter conclusion

This thesis aims to widen the scope of factors used in the economic evaluation of malaria diagnostics and treatments, while seeking to inform policy makers on intervention efficiency on a localised basis. The methods described in this chapter can ensure that evaluations have greater decision relevance in terms of the content, model structures, and how they handle variability and uncertainty. Some of the methods have recently been applied in the context of high income countries and are being applied for the first time in the context of malaria, as described in the following four chapters.

5. Broadening the range of factors in decision models: Non-adherence to diagnostic test results and the potential harm of antimalarials

5.1 Introduction

A central theme of this thesis is that economic evaluation of malaria diagnostics and treatments should aim to encompass the major considerations relevant to decision making. As demonstrated in the literature review in Chapter 2, many previous evaluations have established their results and decision recommendations based on immediate costs and benefits to the patient and provider, obtained from trial results assessing clinical efficacy. A potential limitation of such evaluations is the existence of factors external to clinical trial settings that can play a significant role in determining whether adopting the intervention under consideration is a prudent choice.

The need to expand analyses beyond immediate costs and consequences is particularly acute in the context of malaria management in SSA. A simplistic, linear portrayal of **infection → fever → diagnosis→ treatment → cure** that underlies many evaluations ignores the pervasive and multidimensional nature of malaria and the circumstances in which its management takes place. Such models will fail to address critical factors along this continuum, including the following. First, fever that might not be a product of malaria infection, whether parasitaemia is present or not (Mwangi, Ross et al. 2005). On the other hand, ongoing, asymptomatic parasitaemia can be the cause of other less apparent detrimental health outcomes (Al Serouri, Grantham-McGregor et al. 2000; Breman 2001).

Secondly, when diagnosing malaria, clinicians might not act in accordance with guidelines or diagnostic tests, and often over-diagnose malaria (Amexo, Tolhurst et al. 2004).

Thirdly, treatment strategies can have a broader impact on transmission intensity and the emergence of resistance to antimalarials, with costs and consequences pertaining to society at large (Bloland, Kachur et al. 2003).

Fourth, 'cure' is subject to interpretation, and can relate either to clearance of symptoms or of parasitaemia. Recrudescence and re-infection are routine occurrences, hindering attempts to classify treatments as single events ending in either success or failure. The time to next infection, for instance, might also influence the choice of antimalarials (Dorsey, Njama et al. 2002).

A fundamental requirement in evaluating diagnostics and treatments of malaria is, therefore, to identify and capture factors that are likely to influence their efficiency, many of which are not necessarily observable in clinical trials. As described in the literature review, the incorporation of possible development of resistance has been realized in several of previous economic evaluations of malaria treatments, and has recently also been included in evaluations in the context of antibiotics (Smith, Yago et al. 2005).

In the context of diagnostics, however, economic evaluations have been far more modest in the range of factors included. The evaluation of malaria diagnostic tests has been limited to outcomes pertaining to the immediate patient, and often using proxy measures such as *patient correctly diagnosed* as opposed to estimating their impact on final health outcome. This area in particular was therefore identified for research in the thesis.

The objective of this chapter is to present an evaluation of RDTs compared to microscopy and presumptive treatment that accounts for two previously ignored factors, the non-adherence to the results of diagnostic tests in routine practice, and the possible negative externalities associated with the use of antimalarials. The intention is to explore the impact their inclusion has on the cost-effectiveness of diagnostic tests, and the implications of ignoring them in terms of health outcomes and misallocation of scarce resources.

5.2 *Choice of diagnostic test*

The increasing variety of diagnostic tests available for some of the most widespread diseases in sub-Saharan Africa, including HIV, tuberculosis and malaria, could potentially improve targeting of drugs to those patients most in need. Improved targeting of antimalarial drugs in Africa is an increasingly urgent priority as new ACTs are deployed, costing significantly more than previous and now generally ineffective antimalarials.

The widespread practice of treating any non-specific febrile illness as malaria threatens the sustainability of ACT deployment; in many settings most antimalarials prescribed go to those with no malaria parasites (Amexo, Tolhurst et al. 2004; Hamer, Ndhlovu et al. 2007; Reyburn, Mbakilwa et al. 2007). This means that in addition to the unnecessary use of scarce resources, other potentially severe causes of febrile illness are ignored. Moreover, it exposes patients to potentially toxic reactions to the drugs and is likely to speed up the onset of resistance to ACTs, as large numbers of individuals are exposed to parasites in sub-therapeutic drug levels (Bloland, Ettling et al. 2000).

Reflecting these concerns, WHO and a number of national guidelines in SSA now recommend that treatment for non-severe malaria should, at least for older children and non-pregnant adults, be restricted to those with a positive parasitological test for malaria (NMCP 2005; WHO 2006b; Zurovac, Njogu et al. 2008). In many settings this is difficult to achieve with current blood slide testing that is time consuming, often inaccurate and only available in larger health facilities where a minority of patients seek care (Hetzel, Iteba et al. 2007). Recent improvements in RDTs address many of these problems and their current cost is less than most courses of ACT (Bell, Wongsrichanalai et al. 2006). Several studies have explored the economic consequences of deploying RDTs alongside ACTs and suggest they are cost-effective, but critically assume prescribers respond to negative test results by not prescribing an ACT(Rafael, Taylor et al. 2006; Rolland, Checchi et al. 2006; Lubell, Reyburn et al. 2007; WHO 2007).

5.3 The extent and reasons for non-adherence to the results of RDTs

While these diagnostics might appear to be cost-effective in standard models which assume clinicians act on the results, their deployment could consume scarce funding with little effect if non-adherence to test results is not accounted for.

The assumption that prescribers will adhere to test results is highly questionable. In Zambia where RDTs were used routinely in settings where microscopy was unavailable, antimalarials continued to be prescribed to over a third of patients with negative test results (Hamer, Ndhlovu et al. 2007). In Tanzania, a recent randomised trial of RDTs compared to blood slide testing found that in low transmission areas, over 90% of all antimalarials prescribed were for patients with a negative test result, irrespective of the test method used (Reyburn, Mbakilwa et al. 2007). This very high level of overdiagnosis extended also to patients with severe illness (Reyburn, Mbatia et al. 2004).
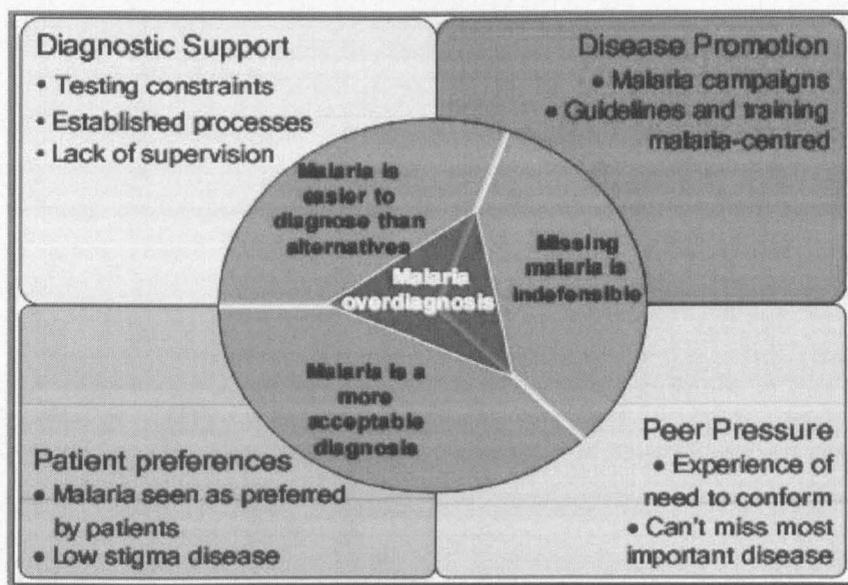
Figure 5-1: Mindline model for the overdiagnosis of malaria. Taken from Chandler et al. 2008. The outer quarters show the influencing factors and the inner circles the 'mind-lines' leading to overdiagnosis

The reasons for ignoring both RDT and microscopy negative test results have recently been explored by Chandler et al., who identified a number of factors that contribute to the practice of overdiagnosis of malaria (Chandler, Jones et al. 2008). The authors identify firstly the spheres of influence on clinicians, such as training programmes, patient

expectations and peer-pressure, followed by the *mind-lines*, meaning the thought patterns that are assumed to lead to overdiagnosis. These consist of the ease and acceptability of diagnosing malaria, and the fact that missing a case of malaria appears indefensible in face of both peers and patients and particularly so given the extensive emphasis placed on treating malaria through promotion and training campaigns. Both the influences and the *mind-lines* are illustrated in Figure 5-1.

5.4 Harm of Treatment

When considering the costs and consequences of diagnosis and treatment, most previous evaluations have addressed only those pertaining to the patient and provider, stopping short of including the broader benefits to society associated with the reduction in antimalarials prescribed, even though these are one of the primary incentives for the introduction of RDTs (Bloland, Kachur et al. 2003; Nosten and Ashley 2004).

As with all drugs, there are a number of possible adverse outcomes surrounding the use of antimalarials. Firstly is their possible toxicity, particularly amongst children and pregnant women (Price, van Vugt et al. 1999; Johann-Liang and Albrecht 2003). Second is the contribution to the risk that this class of drugs also succumbs to parasite resistance (Bloland, Kachur et al. 2003). A small number of economic evaluations have incorporated the development of resistance in determining the cost-effectiveness of antimalarials (Schapira, Beales et al. 1993; Goodman, Coleman et al. 2001a; Coleman, Morel et al. 2004; Yeung 2006), but none have done so in the context of diagnostic tests.

Incorporating this factor in the analysis requires a direct comparison of treatment consequences for the individual patient with those to society more broadly. This in turn requires a shared measure to assess these different consequences on a single scale. In a unique attempt to do so, Rafael et al. (2006), based on a method presented by Girosi et al. (2006), estimated the number of antimalarials used that would eventually lead to the loss of a life though adverse consequences of treatment. They concluded that the use of 200 antimalarials would lead to the loss of a single statistical life (Girosi Rafael et al. 2006;

Rafael, Taylor et al. 2006). Thus the benefits associated with provision of 200 ACTs to patients whom the clinicians suspect to have malaria is offset against the future loss of a single life. This value is supposed to reflect lives lost due to emergent resistance to the drug, possible adverse effects, and the opportunity cost in terms of the resources spent on the treatment and the cost of introducing new regimens should it fail. Rafael et al. then made use of this value in evaluating the number of lives saved with the use of new diagnostic tests that reduced the unnecessary use of antimalarials. The authors did not proceed to include an economic component to their analysis.

The method used to obtain this value made use of observations of previous decisions made by the medical community regarding diagnostic practices. These were assumed to reveal preferences concerning the relative significance of increasing the probability of detecting true malaria cases as opposed to reducing the unnecessary use of antimalarials and the likelihood of correctly diagnosing other non malarial illnesses. The choice of one test over another essentially expresses the trade-off between sensitivity and specificity; this trade-off reveals the benefit that is assumed to be gained by a certain increase in sensitivity, with a higher proportion of true positive cases being found, as opposed to the increase in false positives and unnecessary treatments being used (Girosi et al. 2006). By quantifying the degree to which the medical community was willing to accept a lower sensitivity in favour of an increase in specificity, an estimate was derived for the number of treatments that are assumed to result in a loss of life at some point in the future.

The method, however, has numerous limitations and the authors identify the area as one for further research. First, it assumes that the medical community is tacitly aware of the potential long term costs of unnecessary use of antimalarials and is able to quantitavely account for this in the trade-off with ensuring that true malaria cases are adequately treated. In the context of this thesis use of this method has further limitations – first, there was no breakdown into the different components of the harm of treatment, and the opportunity cost component is already accounted for in the evaluation as captured in the drug price. Secondly, the loss of life associated with the harm of treatment is assumed to occur at some point in the future, although the timing for this is not specified. This implies

that only partial discounting can be applied – that for the subsequent years from a given point of death, but not for the time until that death occurs which is unknown. The value of 200 antimalarials used leading to the loss of a statistical life is therefore likely to be an overestimate and warrants substantial exploration in sensitivity analyses, as shown in the analyses in this thesis.

Despite these limitations, presently this is the only available estimate for the broader costs of using antimalarials. In effect, the inclusion of a parameter reflecting these adverse outcomes in the analysis places a necessary constraint on the otherwise unrestricted use of antimalarials.

The decisions with respect to diagnostics policy makers face are therefore as follows. Firstly, they must decide whether to continue with a strategy of presumptive treatment, or whether to opt for use of diagnostic tests prior to provision of ACTs. Secondly, if opting for parasitological confirmation, they must consider whether to rely on the use of microscopy where this is available, or adopt the use of RDTs in their place.

The following analysis thus compares the costs and consequences of use of RDTs, microscopy and presumptive treatment, allowing for variation in the level of adherence to diagnostic tests and the indirect harm associated with the use of antimalarials. Both diagnostic methods and presumptive treatments are evaluated as these are all likely to continue to play a significant role in the management of febrile patients in the foreseeable future.

5.5 Methods

5.5.1 Data

Data for the analysis were obtained from a randomized control trial carried out in three hospitals in northeast Tanzania in 2005. These hospitals serve areas where the transmission of malaria has previously been characterised as very low, low and high, with parasite prevalence of 2%, 5% and 61% respectively, in febrile children under the age of 5 years presenting at health facilities (Drakeley, Carneiro et al. 2005).

Patients (n 2416) for whom the clinician had requested a parasitological test for malaria were randomised to diagnosis using routine microscopy or an RDT for the detection of Pf Histidine Rich Protein 2 antigen (Paracheck-Pf®). In both arms, reference slides were taken and later double-read according to research methods to determine diagnostic accuracy of the tests. Data on treatments given were recorded and used to compare clinician adherence to microscopy results against those for RDTs; the trial is described in further detail elsewhere (Reyburn, Mbakilwa et al. 2007). The relevant data for the analysis in this chapter are prevalence of parasitaemia amongst febrile patients, clinician adherence to test results, test accuracies (sensitivity and specificity), and costs of both diagnosis and treatment. Costs were obtained from data collected retrospectively in the low and high prevalence settings, supplemented where necessary by data from the literature.

The trial was carried out just before Tanzania rolled out ACTs for first line treatment of malaria in health facilities, therefore patients treated for malaria still received standard monotherapies (most commonly sulfadoxine-pyrimethamine). In order to ensure policy relevance, the analysis here assumed the treatment given was artemether-lumefantrine, which has since been introduced as first line treatment in Tanzania and other countries in the region.

In August and September 2005, the hospitals where the trial was carried out were visited to assess the providers' cost of testing and treating patients with malaria suspected illness. During this time laboratory staff were observed and interviewed to obtain the resources

and time they spent on malaria microscopy. Slide costs were obtained by combining this micro-costing (noting all resources used at the point of delivery, and using an ingredients approach to calculate their total economic cost) with step-down costing of hospital expenditure to estimate indirect laboratory costs. Data for step-down costing were obtained from two of the hospitals' accounts and recent independent evaluations of all their assets. The cost of utilities and service departments were apportioned using measures that reflected their nature, for instance electricity and cleaning services were apportioned according to each direct service department's surface area, while transport and sustenance costs were apportioned according to staff numbers. These costs were then apportioned to the various laboratory tests according to their proportional activity, based on staff estimates for the amount of time spent on each type of test, and the number of tests carried out each month as found in the laboratory records.

The cost of RDTs was obtained directly from the manufacturer (Orchid Biomedical Systems, Goa, India) and included shipment costs plus 10% for local transport and storage. The time required for administration of the tests was assumed to be equivalent to the preparation of a blood slide, based on the opinion of laboratory staff who administered both.

The cost of ACT in 2006 was estimated at $1.6 for an adult course of artemether-lumefantrine (AL), the price negotiated between the World Health Organization and Novartis, the manufacturer (WHO 2006a), and adjusted for patient age. These costs were used in the analysis in place of costs of currently used antimalarials to simulate the switch to AL (Coartem®) in Tanzania as the first line drug for the treatment of uncomplicated malaria.

The cost of treatment for patients diagnosed as malaria negative was estimated from data on the treatment cost of trial patients who received an antibiotic but not an antimalarial. The geometric mean was used for these as they were highly skewed to the right. The costing perspective was that of the provider. Costs were collected in Tanzanian Shillings of 2005 and converted to US dollars ($1=1167Tzs for 2005). The cost for clinician consultation time was not included as this was found to be similar for patients in all trial arms.

Parameter	Estimate used	Source	Comment
Costs:			Costs collected in 2005 Tanzanian Shillings and converted to USD ($1=1167Tzs)
RDT	$0.8	Primary	
Microscopy	$0.28	Primary	Low cost partly result of short reading time
ACT	$1.6 adult dose,	(WHO 2006a)	Quantities adjusted for younger age groups
Antibiotics	$0.42	Primary	Test negatives that were adhered to were assumed to receive drug of this cost
False negative	Varies by age, prevalence with respect to probability of untreated malaria becoming severe and CFRs (see below). Value of YLL=$150	(WHO 2006c)	Value of YLL based on WHO benchmark for 'attractive' interventions
False positive	Determined by proportion of NMFIs that are bacterial, the probability they become severe, and CFRs, Value of YLL=$150		
RDT sensitivity	93%	Primary	
RDT specificity	96%	Primary	
Mic. sensitivity	73%	Primary	
Mic. specificity	93%	Primary	

Probability untreated malaria becomes severe	Under 5 year old	5 to 14 year old	15 years old and above	Source
1% prevalence	0.075	0.050	0.011	
10%	0.075	0.026	0.009	
20%	0.075	0.011	0.006	
30%	0.075	0.010	0.004	Initial estimates based on previous analyses (Goodman et al. 2001; Goodman et al. 2006), then classified by transmission intensity using expert opinion (Christopher Whitty, Hugh Reyburn)
40%	0.075	0.010	0.003	
50%	0.075	0.010	0.002	
60%	0.075	0.010	0.002	
70%	0.075	0.010	0.001	
80%	0.075	0.010	0.001	
90%	0.075	0.010	0.001	
CFR treated severe malaria	0.05	0.10	0.20	
CFR Untreated severe malaria	0.30	0.40	0.50	
Probability NMFI becomes severe	0.01	0.005	0.010	
CFR NMFI	0.1	0.20	0.30	

Table 5-1: Parameter inputs. CFR -Case Fatality Rate; NMFI - Non Malarial Febrile Illness; YLL – Year of Life Lost

Prevalence was defined as the proportion of febrile patients presenting with parasitaemia, and varies with malaria transmission (Marsh and Snow 1999). Prevalence of under 10% was classified as likely in low transmission, 10%-50% in moderate, and above 50% in high transmission settings. Adherence is defined as the percentage of test negatives that did not receive antimalarials, in other words 0% implies all patients with a negative test results receive antimalarials, while 100% implies that none of them do. Non-adherence to positive test results was not examined as this was an extremely rare occurrence (<1%). Prescription of antibiotics to patients with negative test results was as documented in the trial, i.e. just over 50%, independently of provision of an antimalarial. The estimates for test accuracies for both RDTs and microscopy were derived by combining data from all three trial sites.

The parameter estimates used in the analysis and the sources for these are summarised in Table 5-1.

5.5.2 Evaluation framework

As explained in Chapter 4, the analysis uses a cost-benefit framework, accounting for both provider costs, and the monetary value of years of life lost (YLLs) due to incorrect diagnosis and inappropriate treatment. This framework allows for a more coherent analysis in directly observing the impact of compromised adherence to test results and the harm of treatment factor on total costs, without the need for use of cost-effectiveness ratios. As both costs and consequences of the different strategies are expressed in monetary terms, these are differentiated by referring to either direct costs to describe financial expenditures alone, or to total costs incorporating both financial expenditures and consequences in terms of value of life years lost. The option that incurs the lowest total cost is therefore considered the most efficient.

A decision tree structure was chosen as most adequate to synthesise the data, represent alternative options, and indicate the most efficient outcomes (Figure 5-2). A number of parameters were made to interact, for instance the probability of developing severe malaria responds to prevalence and patient age, as do treatment costs for ACTs. This

provides tailored results that are responsive to patient age and location, rather than averaged results that mask the differences between sub-populations.

Patients for whom the test provided an incorrect result to which the clinician did adhere were classified as either false positives or false negatives. Their costs were estimated with the use of a simple flow chart as illustrated in the lower panel of Figure 5-2, following the model used by Goodman et al. (Goodman, Mutemi et al. 2006) and Coleman et al. (Coleman, Morel et al. 2004). Probabilities with which to populate this model were derived from expert opinion as detailed in Table 5-1. The probability of an episode of malaria being self-limiting was determined according to patient age and transmission intensity, with varying case fatality rates according to whether or not the patients were admitted as inpatients.
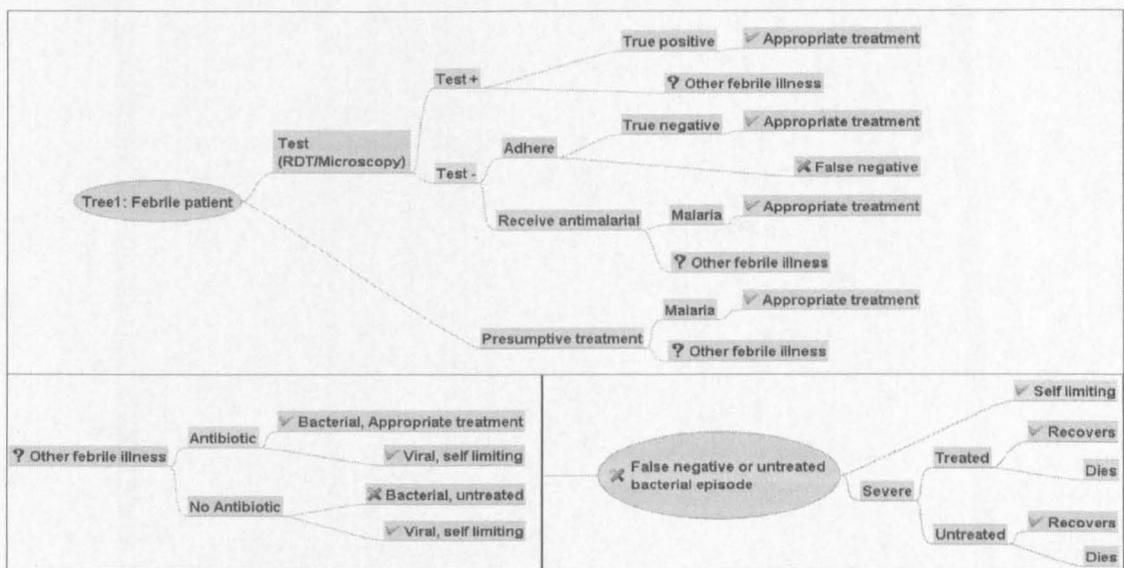
Figure 5-2: Decision trees in the model. The probabilities for developing severe illness and case fatality rates differ with respect to age, transmission intensity, and status as either false negative or untreated bacterial illness, as detailed in table 1.

The death of a patient leads to a number of YLLs, dependent on age and as calculated from relevant life expectancy tables (WHO Statistical Information System 2007), and discounted at 3% (Gold, Siegel et al. 1996). For the primary analysis, a year of life lost was assigned a cost of $150, reflecting WHO's benchmark for an 'attractive' intervention in terms of cost-

effectiveness (WHO 2006c). All other costs and outcomes accrue instantaneously with no further discounting required. The outcomes of the different branches summarise the total costs of the patient under each strategy. The benefits of diagnosis are integrated in this value (as the averted cost of life years lost), so the intervention with the lowest total cost is the most attractive.

The baseline value used to estimate the harm of treatment was that estimated by Rafael et al. (2006), stating that for every 200 antimalarials given one statistical life is lost.

Results are stratified by three age groups – children under 5 years, children between 5 to 14 years, and adults, aged 15 years and above. In addition to the influence of age on transition probabilities and case fatality rates, the responsiveness of the model to patient age is important for policy considerations, since younger patients are still recommended to be treated presumptively in some settings.

5.5.3 Sensitivity analysis

Results were tested for sensitivity to variation in all parameters, and a tornado graph was produced to identify those with highest influence. The greatest degree of uncertainty surrounds the harm of treatment factor. The impact of this uncertainty was tested by both increasing the number of treatments associated with the loss of a statistical life by an order of magnitude, and also by carrying out a threshold analysis to identify the value at which the decision to use an RDT becomes less efficient than presumptive treatment.

5.6 *Results*

Results are presented here by initially describing the costs and accuracies as found in the trial. This is followed by a description of the total cost for each strategy across all prevalences, without accounting for compromised adherence or the harm of treatment, which are introduced in the subsequent two sections. Last are results of the sensitivity analysis exploring the uncertainty surrounding the most influential parameters.

5.6.1 Costs and accuracies

Microscopy costs. The cost per slide was $0.26. Labour was the largest cost component, though staff took an average of less than 1.5 minutes prior to declaring a slide negative or providing a positive result along with a parasitaemia count.

Rapid diagnostic test costs. RDT costs were $0.81, the most significant component being the test itself ($0.60).

Treatment costs. Geometric mean cost for treatment of patients diagnosed as not having a malarial illness was $0.42, as compared to $1.6 for an adult course of Coartem.

Observed diagnostic accuracy and prevalence. Test specificity was comparably high for both tests – 95% for RDT and 93% for microscopy – but RDTs were substantially more sensitive (93%) than routine microscopy (71%).

Choice of strategy without accounting for non-adherence or the harm of treatment. Without accounting for the impact of non-adherence and the harm of treatment, results confirm previous analyses that suggest that use of RDTs is beneficial for children under five years of age at low transmission, and less so at higher ones where presumptive treatment is the preferred option (Figure 5-3). Microscopy is the least efficient option in all areas above 5% prevalence. For adults both diagnostic tests are equally efficient and far more so than presumptive treatment across all but the highest prevalences.
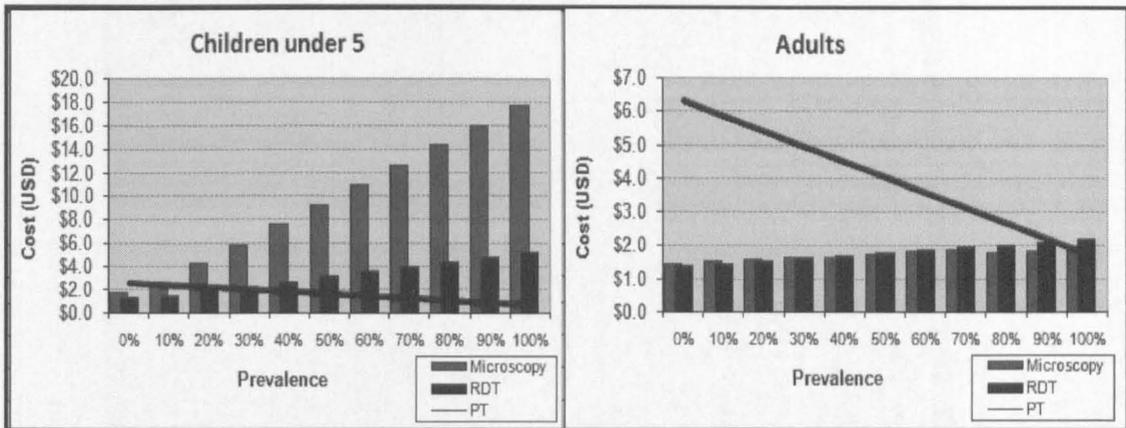
Figure 5-3: Total cost for RDTs, microscopy and presumptive treatment across all prevalences for children under five and adults. PT – presumptive treatment

5.6.2 Adherence

When compromised adherence is allowed for, the model output shows that non-adherence to test results has a significant effect on their cost-effectiveness, and indicates some situations where even modest compromises in adherence to negative test results leads to expenditure higher than that incurred with presumptive treatment. This does depend critically, however, on transmission setting and age.

In a low transmission setting, illustrated by a prevalence of 10%, and for a 15 year old patient, both tests incurred higher costs than presumptive treatment when adherence was below 20%, and application of a test became increasingly attractive as adherence improved (Figure 5-4). For instance at an adherence level of 50%, both tests were less costly than presumptive treatment (as they were below the presumptive treatment threshold), and at this point incurred equal expenditure, approximately $4.6 per patients. Above this level of adherence RDTs became marginally more attractive, but despite their higher sensitivity, they did not have a significant advantage over microscopy in a low transmission setting.

In a high transmission setting, illustrated here by a prevalence of 60%, and for a 15 year old, both tests incurred higher costs than a strategy of presumptive treatment if adherence was below approximately 65% (Figure 5-4). As is evident in the graph, the two tests

followed different trends as adherence improved further. While adhering to RDTs led to substantial cost savings, the cost of microscopy was almost unchanged with improved adherence, due primarily to its low sensitivity, which resulted in significant numbers of false negatives.
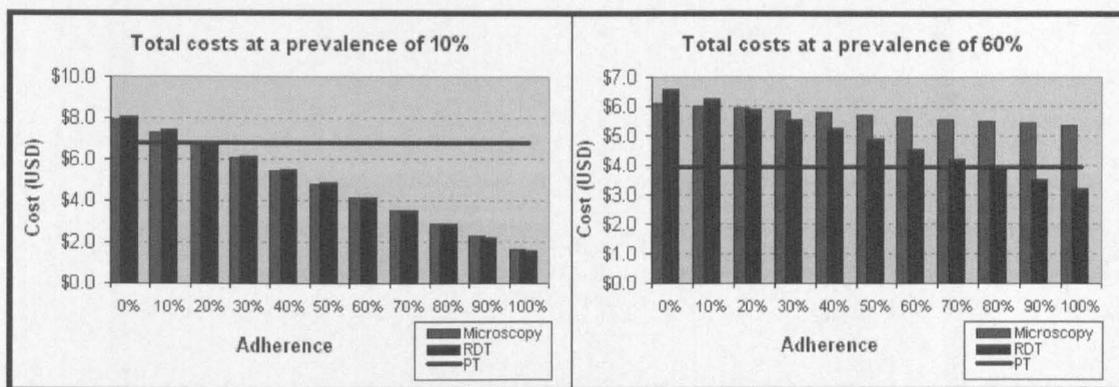
Figure 5-4: Model output for a 15 year old patient, demonstrating total costs for RDTs and microscopy, with varying levels of adherence in low (left) and high (right) transmission intensities. PT- Presumptive treatment.

Figure 5-5 compares each of the tests directly to presumptive treatment for adults across all levels of prevalence and adherence, showing the proportional change in cost when using RDTs and microscopy relative to presumptive treatment. The upper left corners in both charts indicate that in low prevalence settings, use of either test with high adherence to results led to cost savings of over 50% as compared to presumptive treatment. As prevalence increases to the medium-high range however, adherence to test results must increase more than proportionately in order for their use to remain attractive. At very high levels of prevalence both tests appeared more costly irrespective of adherence due primarily to imperfect test sensitivities, and presumptive treatment remained the more efficient option.

109
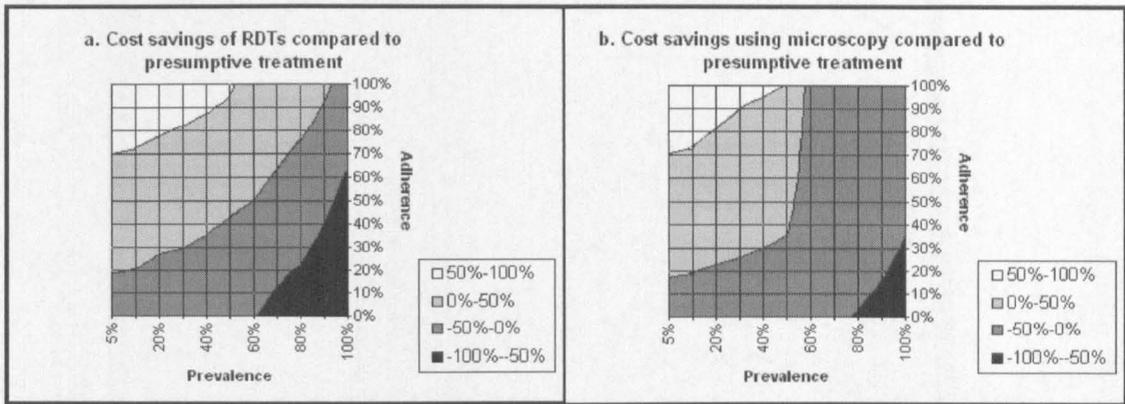
Figure 5-5: Cost savings with RDTs (a) and microscopy (b) for an adult patient using presumptive treatment as a baseline, across all prevalences and levels of adherence. The darker areas indicate that presumptive treatment is more efficient

The level of adherence to RDTs observed in the trial for instance, of approximately 50%, imply that RDTs would increase costs by 62% and 43% in low and high transmission settings, respectively. For microscopy the results were similar at low transmission, while in a moderate to high transmission setting the cost increase was only 10%. At higher prevalence microscopy would become less attractive than presumptive treatment.

Figures 5-6 shows the most attractive strategy at all prevalences and levels of adherence using profiles for patients aged three, seven and twenty five. They indicate that the use of either parasitological test for younger patients was unattractive in settings of medium and high transmission, even if tests were fully adhered to.
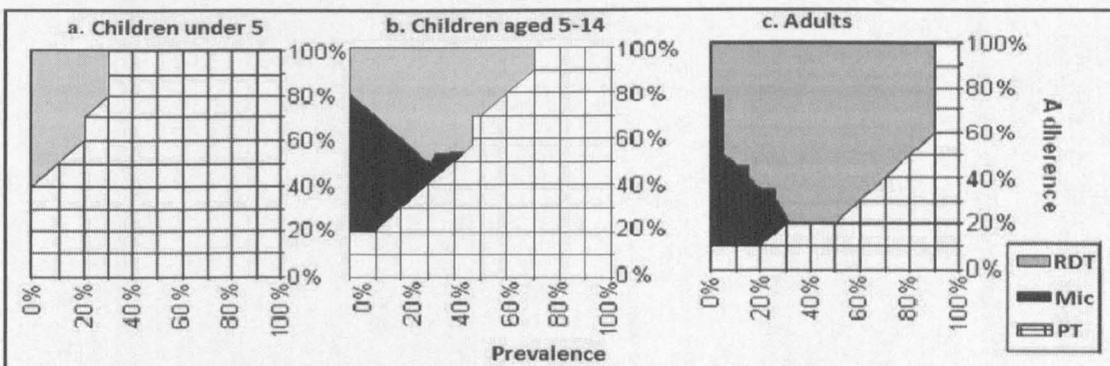
Figure 5-6: Most attractive strategy stratified by patient age. Shading indicates which strategy is preferred at specific combinations of prevalence and adherence. RDT- rapid diagnostic test; Mic- Microscopy; PT- Presumptive treatment

5.6.3 Harm of treatment

Harm of treatment was then included in the model. The baseline estimate for this factor implies that for every 200 ACTs given, one statistical life is lost in the future (Rafael, Taylor et al. 2006). Figure 5-7 shows the difference in total costs with and without the inclusion of the harm of treatment factor for the 5-14 year old age group. Inclusion of the harm of treatment factor has a substantial impact on results, with a considerably higher surface area of the graph indicating RDTs and microscopy being preferred to presumptive treatment.
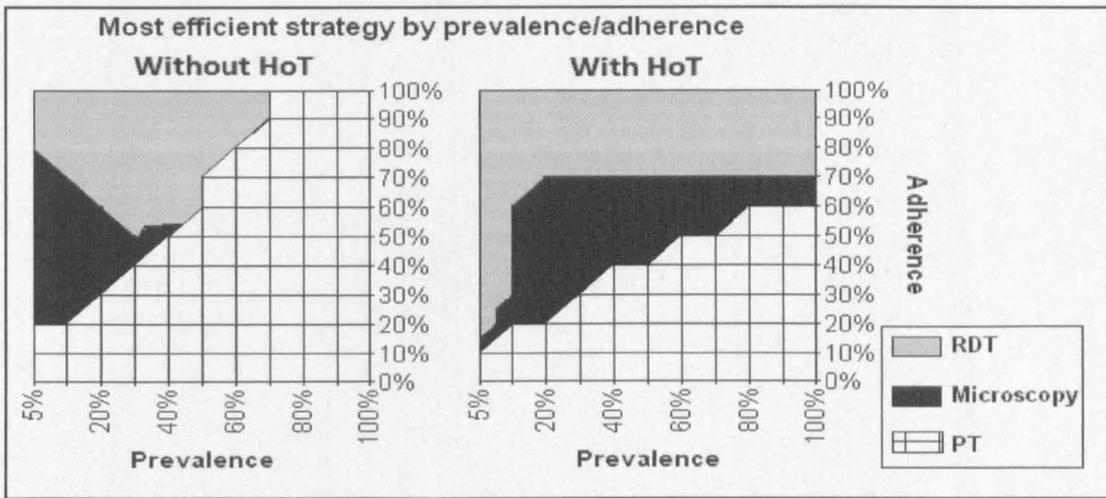
Figure 5-7: Preferred strategy with and without the inclusion of the harm of treatment factor for patients aged 5 to 14. RDT- rapid diagnostic test; Mic- Microscopy; PT- Presumptive treatment

With respect to younger children, without inclusion of the harm of treatment factor, there was almost no advantage to using either RDT or microscopy as compared to presumptive treatment (Figure 5-6a above). When the harm of treatment associated with over-prescription of antimalarials is included, results change considerably in favour of either diagnostic test (Figure 5-8).
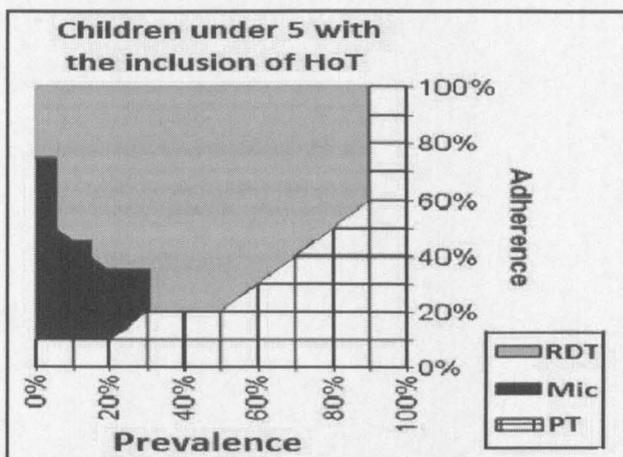
Figure 5-8: Model output for children under 5 with the inclusion on the harm of treatment factor. RDT-rapid diagnostic test; Mic- Microscopy; PT- Presumptive treatment

5.6.4 Sensitivity analysis

Results were most sensitive to the cost of a YLL and variation of the harm of treatment factor. Higher values of YLLs led to scenarios that were more costly and more effective being considered more attractive. If the cost of YLLs is set to zero (i.e. the value of health outcomes is ignored), RDTs would never be the most efficient option, while microscopy was still attractive although decreasingly so as prevalence increased and adherence fell.

Figures 5-9a-c demonstrates the circumstances under which each of the strategies is most attractive, stratified by the value of a YLL averted and for an adult patient. At a YLL value of $25 ('very attractive' (Edejer, Baltussen et al. 2005)), RDTs gained some advantage in the mid prevalence range as long as high levels of adherence were maintained, with microscopy remaining the preferred option for low prevalence areas even at low levels of adherence. Using $150 per year of life lost, RDTs became the preferred option up to a prevalence of about 70%, where presumptive treatment became the more efficient option. At a value of twice the Tanzanian GNI per capita for the year 2005 (World Bank 2007), an alternative rule of thumb (Garber and Phelps 1997), i.e. $680, RDTs dominated across all but the lowest levels of adherence and highest levels of prevalence.

112

a. Value for YLL averted=$25   b. Value for YLL averted=$150   c. Value for YLL averted=$680
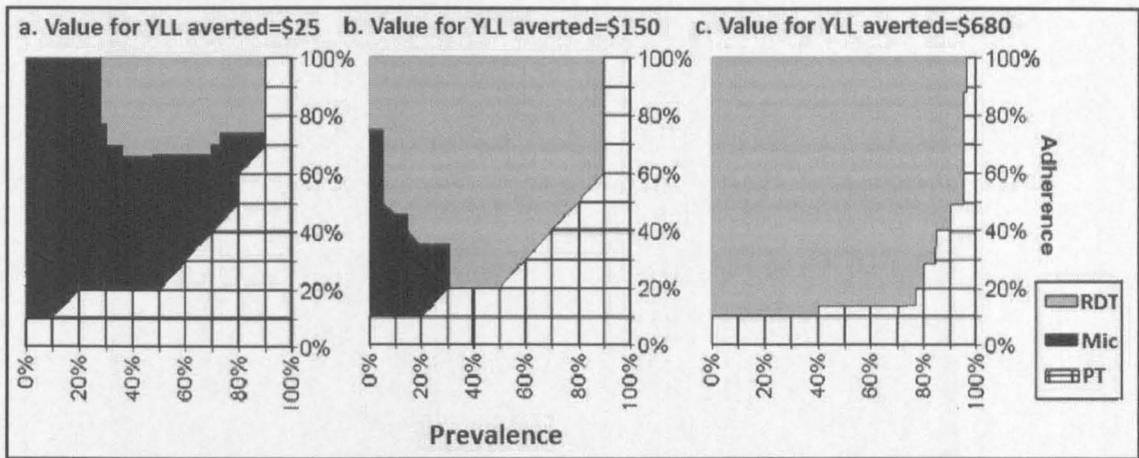
Figure 5-9a,b,c: Most efficient strategy (indicated by shading) by adherence and prevalence, with respect to value of a year of life lost.

Recognizing the uncertainty surrounding the harm of treatment parameter, a second value of 2,000 ACT treatments per death was arbitrarily chosen to observe the sensitivity of results to a lower estimate of harm of treatment (Figure 5-10). Even with this much lower estimate of harm of treatment, both microscopy (not shown) and RDTs remained more efficient than presumptive treatment in prevalences of up to 35%, as long as adherence was high. In fact in an area of medium transmission intensity, the number of ACTs equating to the loss of a statistical life would have to be as low as 7,000 treatments before presumptive treatment becomes the more efficient option.
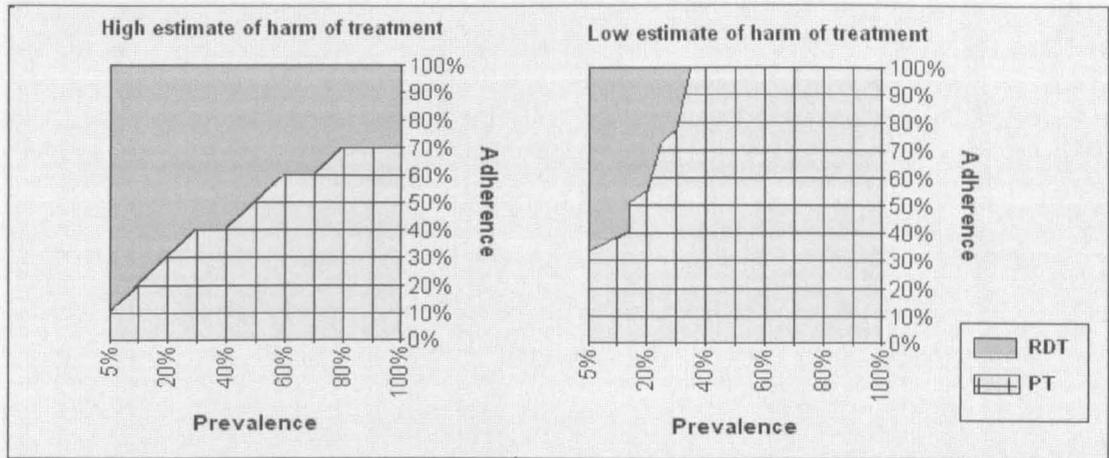
Figure 5-10: Total costs for children under 5 with the baseline estimate of the harm of treatment (left), and a low estimate (right). Most efficient strategy indicated by shading

113

5.7 Discussion

5.7.1 Limitations

A number of potential limitations of the methods should be acknowledged. First, some of the data within the model had to be estimated from indirect sources and expert opinion because locally relevant data do not exist in the literature. Data for case fatality rates for untreated malaria, for instance, cannot be accurately measured, and the interaction between these, transmission intensity and age, adds an additional level of uncertainty. While it is imperative that studies provide explicit statements of the estimates used, models could also be developed that allow users to enter their own estimates to observe the impact of their variation on final results, as is described in detail in Chapter 6. Expert opinion could also be gathered from a wider range of individuals using methods such as a Delphi survey to produce a consensus, a project which is currently underway (Annex 2).

Second, the antimalarial effectiveness in the analysis was assumed to be 100% effective. This is not an entirely accurate representation as AL and other ACTs do exhibit treatment failures, although to a lesser extent than most existing monotherapies (Mutabingwa, Anthony et al. 2005). In this instance the effect of a small proportion of treatment failures was assumed to have a similar impact on the efficiency of all strategies, therefore was excluded from the analysis.

Thirdly, characterising transmission is prone to numerous difficulties and the methods of estimating transmission intensity from hospital data are inevitably not precise. This study used prevalence of parasitaemia amongst febrile patients, although this method is subject to a number of limitations, most significantly that it does not account for individual patients' levels of parasitaemia. This has been suggested to provide a better indication of whether the infection is in fact the cause of illness (Marsh and Snow 1999).

The model was not designed to perform probabilistic sensitivity analyses as the main aim of the study was to demonstrate the variation in results in response to changes in the

parameters of interest (adherence and the harm of treatment, across a range of prevalences). Many of the other parameters in the model would not be suitable for use in a PSA as they are not data driven parameters, but rather are determined by choice of analytical methods (Briggs 2000).

A final limitation of the analysis was that it did not encompass patient costs, but only those of the provider. This arose because the economic analysis was carried out retrospectively and no data were available for patient costs. An *ad-hoc* attempt was made after the end of the trial to measure the time patients spent waiting for microscopy results, and this suggested that patients endured waiting times often exceeding several hours and sometimes had to return the following day. At that point RDTs were no longer being used but it is likely that they can greatly reduce waiting time and productivity losses for patients and their carers.

5.7.2 Implications for policy and future economic evaluation of malaria diagnostic tests

A major barrier to ACT deployment is its high cost. Targeting of drugs to those in greatest need will support programme sustainability. Diagnostic tests themselves have a cost, and have to be cost-effective if they are to be deployed. This study demonstrates that in the trial setting, clinician adherence to test results had a major impact on the cost-effectiveness of both microscopy and RDTs, as did the inclusion of the harm of treatment factor, although the impact of both factors varied with age and transmission intensity.

In the low transmission setting, testing remained attractive even when adherence was relatively poor, while at higher prevalences adherence would have to increase more than proportionately in order for tests to remain attractive. At very high prevalences, however, presumptive treatment remained attractive given the imperfect sensitivity of tests (particularly microscopy) under field conditions, where high adherence combined with poor sensitivity resulted in increased costs.

Policy makers could consider investing in training programmes for clinicians to encourage adherence to RDTs, prior to their widespread deployment. One such programme in Zambia has shown very positive results (Harvey et al. 2008). Cost assessments of such programmes could be combined with the results from this analysis to estimate their own cost-effectiveness.

Regarding the inclusion of the harm of treatment factor, results of this study indicated that particularly amongst young children this could spell the difference between continuing with presumptive treatment and embracing a strategy of diagnostic confirmation. This population is of particular significance as this is where the majority of malaria cases occur. The parameter however is surrounded by considerable uncertainty. The only estimate available in the literature employed methods that have significant limitations, as acknowledged by the authors (Girosi et al. 2006). More precise quantification of this is difficult, for example assessing the relationship of quantities of ACTs used and patient compliance with their proper use to the development of resistance. Alternative possible methods to quantify this parameter are the use of Delphi surveys to obtain a wide variety of expert opinions in an explicit manner, or alternatively estimating the quantities of particular antimalarial usage in specified regions and periods, and assessments of the subsequent deaths due to subsequent treatment failures due to parasite resistance.

In this case the parameter was varied by one order of magnitude, followed by a threshold analysis to determine the point where presumptive treatment again became more efficient. Across a wide range of values the decision recommendation favoured the use of RDTs.

Excluding the harm of treatment parameter would essentially equate to stating that there are no long term costs associated with widespread use of antibiotics or antimalarials, an assertion that has proved erroneous and exerted high costs in human life with respect to the failings of chloroquine and other antimalarials.

Decision models are often constructed using a relatively limited scope of parameters to ensure that all parameter values used are highly defensible. This can mean that only those

factors that are readily observable and measurable, ideally in an RCT, and attain statistical significance, are included in the analysis. Many RCTs, for instance, now routinely collect cost as well as effectiveness data prospectively to obtain more robust estimates for subsequent cost-effectiveness analyses[7] (Briggs 2000). However even where direct costs are adequately captured in a clinical trial, there remains a range of other potentially decisive factors commonly excluded from evaluations, for example:

- Factors that are characteristic of routine practice but not of clinical trials, such as compromised adherence to guidelines

- Long term health outcomes for patients

- Externalities (i.e. intervention costs and consequences that might not relate directly or uniquely to the patients or providers, but to carers, families or society at large)

Capturing these issues in economic evaluations introduces levels of uncertainty that are incompatible with standard statistical paradigms that either accept or reject a hypothesis at certain arbitrary thresholds. The reason these parameters are excluded is that it is difficult to estimate them with a degree of certainty exceeding widely acceptable p values, such as the 0.05 threshold. In a more practical decision making context, where factors such as these can play an influential role in determining the cost-effectiveness of an intervention, it is imperative that these are included in analyses albeit with thorough exploration of the uncertainty surrounding them.

---

[7] As Briggs points out this often requires a larger sample size, as cost data tends to exhibit high variance and can be heavily skewed; this raises potential ethical concerns, as difference in treatment effectiveness might already be determined, yet additional patients will continue to receive the inferior treatment to reach the necessary sample size for cost data.

5.8 *Chapter conclusion*

An increasing variety of tests are being made available for routine diagnosis of AIDS, malaria, tuberculosis and other common diseases. The cost implications of non-adherence to test results have largely been ignored in previous evaluations, despite the overwhelming evidence that it is a problem throughout Africa. This analysis has demonstrated the importance of allowing for compromised adherence, both in terms of its implications for expenditure on diagnostics and treatment and for health outcomes. It also provides a foundation for estimating how much policy makers might consider spending on programmes to improve adherence to test results.

The inclusion of the harm of treatment factor incorporated a second essential element into the evaluation of diagnostic strategies, relating to the benefits of reducing unnecessary use of antimalarials. Despite considerable uncertainty around this factor it remains one of the primary reasons for considering the use of RDTs and other diagnostic tests and its exclusion renders evaluations less relevant for decision making purposes. In modelling the costs and consequences of malaria diagnostics and treatments, it is imperative that all major factors are accounted for prior to their widespread deployment, as failing to do so can lead to inefficient decision-making in areas of the world that can least afford it.

6. The localisation of economic evaluations using Decision Support Tools – comparison of methods

The aim of this thesis is to develop decision models that incorporate a broader range of factors than immediate costs and effectiveness into the analysis, factors which should be accounted for when considering the adoption of malaria diagnostics and treatments. The thesis also aims to ensure that decision models account for local variations in factors such as malaria epidemiology and antimalarial effectiveness. The previous chapter broadened factors included in the economic evaluation of malaria diagnostics and treatments. This chapter focuses on increasing the relevance of decision models to different settings. In the chapter, two alternative models are developed as decision support tools for choosing between competing RDTs. These models provides policy makers with the ability to explore the decision options they face, the uncertainties involved and the immediate and long term outcomes for patients, providers or society as a whole. Two alternative DST designs are developed in order to explore the trade-off between complexity and methodological validity.

6.1 Introduction

Standard economic evaluations make use of decision models that facilitate the convergence of two criteria – costs and effectiveness. These models are then used to generate results which can be tested for generalisability to other settings; the study might then be published declaring one intervention more or less cost-effective than its comparator. The data input into evaluations, however, often include highly variable parameters such as intervention costs, effectiveness and target population characteristics. Variation can be distinguished from parameter uncertainty, where the parameter value is unknown (Briggs 2000). Variability poses an impediment to the generalisation of evaluation conclusions (Bryan and Brown 1998). While sensitivity analyses can allow for some testing of the robustness of results to change in individual parameters, when several

parameters vary simultaneously, the ability of policy makers to assess the relevance of the study to their own settings becomes questionable.

In the context of malaria diagnostics for instance, transmission intensity can vary widely within small areas (Ye, Kyobutungi et al. 2007), with significant implications for the predictive value of diagnostic tests. Input parameters can also vary considerably over relatively short periods of time, as has been the case with effectiveness and costs of antimalarials. Evaluations of malaria diagnostics and treatments published as recently as December 2007 were based on costs of ACT that are now 70% lower (Lubell, Reyburn et al. 2007; Shillcutt, Morel et al. 2007). Antimalarial efficacy has also varied over relatively short periods, sometimes dropping below levels acceptable by WHO standards in a matter of a few years from their introduction, as discussed in Chapters 1 and 4. Given the inevitable time lag between collection of data and the publication of evaluations, it should be recognized that evaluation findings may be out of date, although this is rarely acknowledged in, for instance, studies that cite previous findings. Incorporation of most recently available data in evaluations can help ensure that better decisions are made, avoiding high human and financial costs due to inappropriate policy options being pursued. This is of particular significance in low income countries where the pursuit of inappropriate strategies is least affordable.

In addition to variation in data inputs, local decision makers might also have priorities and preferences different to those of the analyst regarding which parameters should be included in analyses and the values used. For example, an analysis that was carried out taking a societal perspective might be less relevant for a decision maker interested in the immediate direct expenditure of the health care system. While analysts might have strong beliefs as to which perspective is most appropriate, there is also a strong argument in favour of allowing decision makers to dictate which perspective is appropriate (Weinstein, O'Brien et al. 2003). Similarly, analysts and decision makers might have different interpretations of equity and the choice of target population can have significant impact on an intervention's viability. While an analyst might carry out the intervention under the assumption that all patients should have access to the intervention for the sake of equity,

a decision maker might be interested in exploring how costs and effectiveness might change when intervention availability is limited to a particular sub-groups.

An additional reason why standard evaluations are in some instances less relevant to decision makers' needs relates to the diffusion of decision making powers away from international and national bodies. The decentralisation of health care services has been one of the central reforms characterising health systems in both developed and developing countries (Bossert 1998). While this has taken on different forms in different systems, it has generally led to a diffusion of decision making powers, encouraging decision makers to adopt interventions that are better suited to their particular settings than ones that might have been centrally dictated. Economic evaluations however often do not reflect this reality, in that they continue to seek generalisable conclusions irrespective of local variability.

Whether or not evaluation conclusions do maintain their relevance to other settings, the mere existence of such variability can undermine the confidence that policy makers may have in using the evaluation for their own settings. Without the ability to systematically modify an evaluation to their setting, a decision maker faces one of two options, either to accept the results with only minimal ability to question the relevance of the data to her own setting, or to reject the conclusions as irrelevant. Both options can result in poor decisions being taken.

Alternatively, evaluations can be adapted for use as decision support tool (DSTs), where analysts make use of advancements in IT to make their models available and adaptable for stakeholders to use in their own settings. This is being increasingly recognized as an alternative to trying to generalise evaluation outcomes beyond what is supported by the data (Briggs 2000; Cooper, Sutton et al. 2007; van Gool, Gallego et al. 2007, Scott and English 2008).

6.2 Overview of DSTs in health care

There are very few instances where DSTs have been made available to stakeholders for adaptation of economic evaluations to their own settings. In the context of malaria, for instance, only one such model was identified (Shillcutt, Morel et al. 2007), which appears to be more than in other fields of health[8]. One area where such models can be found in greater abundance is the context of decision support for patients and clinicians in choosing between treatment options. In this context the models aim to converge a range of criteria for each option (e.g. effectiveness, side-effects, etc.), often by weighting them by significance for the patient, and systematically valuing each criterion.

While there are a number of proposed model frameworks for this, and a large body of literature on the theory behind them, there is a large gap between the theoretical literature on decision support tools and their application in practice. A review of the instances where Analytical Hierarchical Process models have been used (one of the most widespread forms of DSTs), reveals that in total these have been documented in use approximately 50 times, almost half of which were in the context of individual patient care options (Liberatore and Nydick, In press). In those instances where the decisions involved resource allocation, this was done in the context of highly specific issues such as a model to inform hospitals on choice of ventilators to purchase, without concern for their cost-effectiveness relative to other interventions (Chatburn and Primiano 2001). Many of these analyses were also carried out for model testing purposes rather than routine use.

The use of DSTs for resource allocation is, therefore, a relatively new approach to the conduct of economic evaluations. While the model structures required for DSTs are on the whole similar to those of standard evaluations, DSTs potentially serve as a powerful tool for policy makers and other stakeholders to actively explore policy options with greater

---

[8] Based on a rapid review of the literature on decision modelling, and the opinion of a number of other health economists at LSHTM.

confidence that the analysis is relevant to their own circumstances and priorities. In the following section two methodologies that can be used for constructing DSTs are reviewed.

6.3 DSTs – choice of methodological approach

Decision models in standard economic evaluations can be developed broadly as either deterministic, stochastic, or fully Bayesian, differing primarily in how data are interpreted and represented in parameter estimates and in the treatment of the uncertainties surrounding these. While this is true for decision models in general, it has particular bearing on how DSTs are designed in terms of the input of local data and policy makers' preferences and priorities.

Data obtained in local trials and studies can be interpreted in one of two ways. A classical approach aims to test the parameters individually for statistical significance to ascertain how representative they are of the 'true' parameter value. The data can then be incorporated as point estimates in a deterministic model (Drummond, O'Brien et al. 2005) or as probability distributions in a stochastic one. Either way the parameter estimates rely entirely on the given data. A Bayesian approach, on the other hand, takes a broader perspective and views the data in light of previous evidence and beliefs. In so doing, it treats the data as random points from a probability distribution representing the possible parameter values. This distribution can be ascertained by starting with a relevant prior distribution representing the initial estimates of parameter values and updating this with local data to derive the posterior distribution. The priors might be based on existing data from other sites, or derived from expert opinion (Spiegelhalter, Abrams et al. 2004).

With respect to uncertainty, deterministic models will vary parameter estimates individually (or at most 2 or 3 at a time) to observe how the variation influences results. Stochastic and Bayesian models, on the other hand, use probability distributions to represent the uncertainty surrounding an estimate, and then use sampling methods to draw on these simultaneously. Processes such as Markov Chain Monte Carlo simulations can then be run to incorporate all parameter uncertainty and prior estimates of their

values to allow for a single output with a measure of certainty attached (Spiegelhalter and Best 2003).

The remainder of this chapter explores the strengths and weaknesses of these approaches by developing and comparing a deterministic and Bayesian locally adaptable decision support tool for the evaluation of strategies for the diagnosis of malaria. The models are compared in terms of structural differences, the nature of their output, and how decision makers might interact with them when considering policy options. In the next section a brief background to the decision problem is given. This is followed by a description of the two models and the results they provide[9]. The advantages of using a stochastic model without the use of priors are captured in the Bayesian approach and therefore a stochastic model is not presented independently, although the use of such a model is recognized as a valid option. The final section compares the advantages and drawbacks of the two approaches.

6.4 Background to the decision problem: Choice of rapid diagnostic tests

With an increasingly large number of RDTs available on the market, decision-makers must consider a number of factors in determining which diagnostic test is likely to be most appropriate in a particular context. Some of these relate to qualities of the RDT itself, such as target antigen, sensitivity, specificity, shelf-life, heat sensitivity and cost. Other factors relate to the demographic and epidemiological circumstances of areas where the tests are to be deployed. Some data are available, for example from field studies of different RDTs' accuracy in various settings, although this has been shown to vary even within a single

---

[9] While the deterministic model was developed for practical use and has been disseminated amongst stakeholders (Lubell et al 2007b; Lubell et al. 2008b), the Bayesian model was developed for illustrative purposes only.

country or region, presenting a complicated picture to decision-makers (Reyburn, Mbakilwa et al. 2007; Hopkins, Bebell et. 2008).

Given the variation in availability, performance and prices of diagnostic tests and treatments over time and location, and in transmission intensity and host immunity, it is unlikely therefore that any one RDT would maintain its advantage indefinitely or across all endemic areas. Similarly, economic evaluations of an RDT carried out in one setting may not apply in others, or may lose their relevance within a relatively short time as epidemiological patterns and the characteristics of competitor tests change. For these reasons, policy makers might benefit from decision aids that incorporate available data and parameter estimates for factors that are variable, to provide up-to-date recommendations for choice of RDT relevant to their circumstances.

6.5 *Factors for consideration in choice of RDT*

The presumptive treatment of fever episodes as malaria results in significant overuse of antimalarials and delays diagnosis of other illnesses (Olivar, Develoux et al. 1991; Chandramohan, Jaffar et al. 2002; Ndyomugyenyi, Magnussen et al. 2007). Therefore, an important potential gain from introducing a new diagnostic test is in reducing the proportion of febrile patients who receive unnecessary antimalarial treatment. This safely reduces the cost of giving unnecessary antimalarials, and may help to avert morbidity associated with untreated non-malaria illness. An ideal RDT should therefore have high specificity to avoid false-positive results that would prompt unnecessary antimalarial treatment. At the same time, it is critical that an RDT must have high sensitivity to ensure that true cases of malaria are detected and treated appropriately.

In reality, improved sensitivity often comes at the expense of reduced specificity, and vice versa; however, it is difficult to weigh the implications of this trade-off for an individual patient or for public health, as they are often not directly comparable (Girosi, Rafael et al. 2006). Mistakenly diagnosing a patient as uninfected (a false negative) may have more serious clinical consequences than mistakenly diagnosing a patient as infected (a false

positive), but this will not always be true. Extensive overuse of antimalarials is also likely to come at a considerable cost over the longer term due to increased drug pressure leading to possible development of drug resistant parasite strains (Bloland, Kachur et al. 2003).

The trade-off in sensitivity and specificity is apparent in the reported accuracies of the two main classes of RDTs which currently appear most suitable for clinical use, detecting either histidine-rich protein-2 (HRP2) or *Plasmodium* lactate dehydrogenase (pLDH). HRP2 based assays have shown good sensitivity in a variety of field settings, and are increasingly advocated where reliable microscopy is not available (Bell 2002; Rafael, Taylor et al. 2006). Their potential disadvantage, however, is the detection of persistent circulating antigen for up to several weeks after parasites have been eradicated, leading to false positive results (Mayxay, Pukrittayakamee et al. 2001; Tjitra, Suprianto et al. 2001; Singh, Saxena et al. 2002). This may limit the usefulness of HRP2-based assays in areas of high malaria transmission. pLDH-based RDTs appear to be less sensitive but are more specific than HRP2 ones, as this antigen is rapidly cleared from the bloodstream (Piper, Lebras et al. 1999; Moody, Hunt-Cooke et al. 2000; Swarthout, Counihan et al. 2007). HRP2 and pLDH based tests also differ in the parasite species they detect: the HRP2 test detects only *Plasmodium falciparum*, while the pLDH test detects all four human malaria species.

For two main reasons, evaluations of diagnostic tests should also account for relevant differences in malaria epidemiology and population characteristics. Firstly, transmission intensity determines prevalence of parasitaemia and therefore, the probability of a test result being correct (the positive and negative predictive values). Secondly, in high transmission areas the population develops partial immunity with age (Snow 2000). An adult in a high transmission area, for instance, is more likely to be parasitaemic, but much less likely to develop severe malaria. A child in a low transmission area, on the other hand, is less likely to be parasitaemic but more likely to develop severe malaria once infected. The implications and benefits of using an RDT in each setting therefore differ (Zurovac, Midia et al. 2006; Lubell, Reyburn et al. 2007; Shillcutt, Morel et al. 2007).

6.6 Modelling the costs and consequences of alternative diagnostic strategies

The economic framework adopted for the models followed that presented in Chapter 4. A monetary value is placed on adverse health outcomes that arise as a consequence of incorrect diagnoses for each test. These are added to the test and treatment costs. The societal costs associated with use of antimalarials are summarised in the 'harm of treatment' factor, as described in Chapter 5, and this is also added to give the total cost of the diagnostic strategy. The total cost for presumptive treatment is also calculated to indicate when use of either test is inefficient.

The most efficient test is then determined by identifying that with the lowest total cost. Figure 6-1 shows the possible patient progression paths and related costs following the use of a diagnostic strategy. Both the deterministic and Bayesian models are structured in accordance with these progression paths.
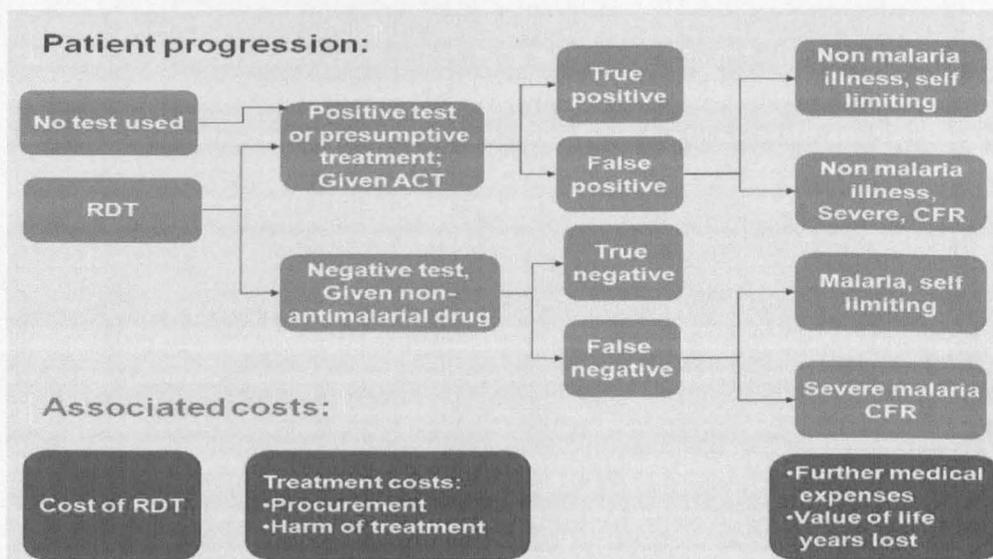
Figure 6-1 Patient progression paths and subsequent costs ; ACT – Artemisinin Combination Therapy, CFR – Case fatality rates, RDT – Rapid Diagnostic Tests

The context for the models is Uganda and the models evaluate the efficiency of two widely available RDTs considered for use in low level health care facilities where microscopy is not available, relative to continued presumptive treatment. Data for these models were

obtained from a Ugandan trial where the RDTs were being evaluated in different sites with a range of transmission intensities (Hopkins, Bebell et al. 2008). Test accuracies were estimated for each of the sites. Costs of RDTs and antimalarials were obtained from the Ugandan Ministry of Health (MoH) in 2007, while inpatient care costs were collected at a single site in southwest Uganda, adjacent to one of the RDT evaluation sites, as described in Chapter 8. Other parameters were obtained from the literature and expert opinion, as detailed in Table 6-1.

Parameter			Base estimate		Source
ACT cost			$1.8 (adult dose)		Uganda MoH
Antibiotic cost			$0.4 (adult dose)		Primary data – Joint Medical Store
RDT 1 cost			$.51		Manufacturer
RDT 2 cost			$.55		Manufacturer
Harm of treatment with ACT or antibiotic			Every 200 ACT or antibiotic doses used result in the loss of one statistical life		(Rafael, Taylor et al. 2006)
Inpatient care severe malaria			$12		Primary data
Inpatient care severe NMFI			$20		Primary data
Year of Life Lost (YLL)			$150, $840		(Sachs 2002; Evans 2004; WHO 2006)
Accuracies Source: Hopkins et al. (2008)	**Site 1**	**Site 2**	**Site 3**	**Site 4**	**Site 5**
RDT1 Sensitivity (number of correct results)	94% (34 / 36)	88% (323 / 365)	76% (339 / 445)	88% (431 / 491)	99% (822 / 832)
RDT1 Specificity (number of correct results)	99.6% (960 / 964)	99% (631 / 635)	99% (547 / 555)	99.6% (507 / 509)	95% (159 / 168)
RDT2 Sensitivity (number of correct results)	100% (47 / 47)	97% (399 / 411)	99% (541 / 547)	98% (566 / 575)	99.9% (874 / 875)
RDT1 Specificity (number of correct results)	81% (770 / 953)	97% (570 / 589)	83% (378 / 453)	95% (403 / 425)	70% (87 / 125)

Table 6-1: Initial parameter estimates used in the models. NMFI – Non malarial febrile illness. ACT – Artemisinin Combination Therapy

6.6.1 The deterministic model

This chapter focuses on the model design, while details of the output are of less relevance and have been published elsewhere (Lubell, Hopkins et al. 2008b). The model was designed using Microso Excel® 2002 and macros were written with Microsoft Visual Basic® 6.3.

A number of functions were added to the standard decision tree structure, making the model adaptable to local circumstances and policy makers' considerations and preferences. Input variability was introduced for parameters that were assumed to vary by locality, and for the particular costs and accuracies of the RDTs under consideration. Users could also choose the age group for which they wished to obtain results, as these are likely to differ for infants, children, and adults. Parameters with high degrees of uncertainty were set initially with best available estimates, which users could modify with their own data and estimates. The model accommodated the possibility that clinicians might continue to prescribe antimalarials in the face of negative test results by allowing for the levels of adherence with negative test results to be altered, as this has been observed to occur with diagnostic tests for malaria (Zurovac, Midia et al. 2006; Reyburn, Mbakilwa et al. 2007)

6.6.1.1 The model interface

All previously mentioned parameters can be varied in the user interface. Other changeable parameters include the probability of developing severe illness by age and transmission intensity, the case fatality rates for malaria and non-malarial febrile illness, and the probability that clinicians adhere to test results.
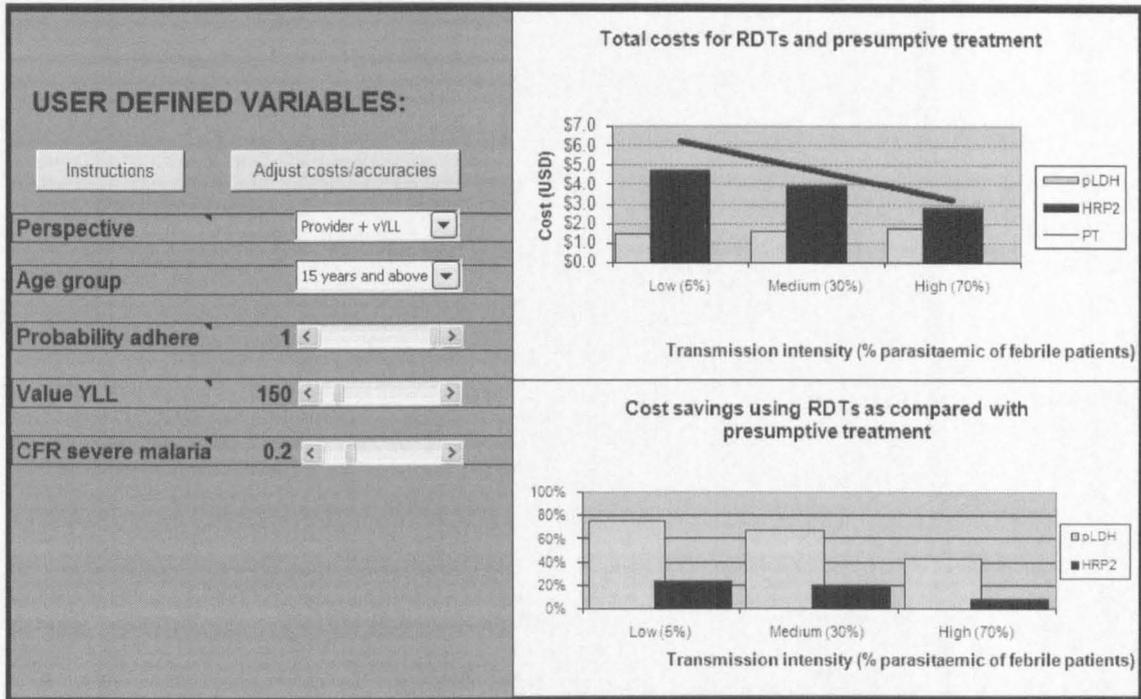
Figure 6-2. The deterministic model user interface. The panels on the left allow variation of input parameters; the button on the right opens a dialogue box where local data on costs and accuracy can be entered. The top right panel depicts the difference between the total cost of each RDT and that of presumptive treatment indicated by the trendline. The bottom right panel shows the proportion of cost savings of each RDT using presumptive treatment as the baseline; CFR – case fatality rate

The user can also choose the perspective of the analysis. Taking the provider financial perspective considers only direct costs of tests and treatment. Alternatively the value of years of life lost to patients due to incorrect diagnosis can be added to the analysis and varied to capture immediate health benefits for the patients. Finally, a societal perspective can be taken, with the incorporation of the harm of treatment factor.

The model output is displayed on two graphs reflecting the difference in total costs in both absolute and relative terms, across three transmission intensities, defined by prevalence of parasitaemia amongst febrile patients (Marsh and Snow 1999). Low transmission was characterized by a prevalence of 3% parasitaemia, medium by 30%, and high by 70%. This allows users to view the most appropriate RDT with respect to regional and seasonal variation in transmission intensity. In the top right panel of Figure 6-2, the trendline represents the total cost in US$ of presumptive treatment in absolute terms, while each

130

set of bars is the cost for either RDT at each transmission intensity. Where the bars fall below the trendline, use of the RDT would, therefore, be more efficient than presumptive treatment. In the lower panel the results are displayed in relative terms, using presumptive treatment as the baseline, so the bars represent the percentage by which RDTs are more efficient than presumptive treatment. Both graphs are included as in some cases the difference in relative terms might seem small, but is large in absolute terms, and vice versa.

The user can examine a variety of 'what if' scenarios, to explore policy options for different sub-populations and settings and also observe how the uncertainty surrounding particular parameters influences results.

When setting the age group to adults for instance, the model suggested that both tests were considerably more efficient than presumptively treating the population, across all transmission intensities. Should the user choose a younger age group, results (not shown) indicated that RDT 1 would hold a strong advantage at low transmission intensity while presumptive treatment would still be the preferred option at higher levels. These results can all be tested for change according to variation in other parameters, such as the probability that clinicians adhere to test results, the value placed on years of life lost, or case fatality rates associated with severe malaria.

The deterministic model therefore enables the user to explore a range of parameter estimates and observe how these influence results. It makes explicit the uncertainty of parameter values and encourages users to provide their own best estimates. If results are visibly robust to changes, this may reduce policy makers' concerns about the uncertainties surrounding decision-making. Where results are sensitive to change, this may encourage consideration of more detailed policies or investment in further research.

6.6.2 The Bayesian model

The Bayesian model was written using WinBUGS 1.4.1. The model was structured in a similar way to the deterministic model, with the additional use of probability distributions to reflect parameter uncertainties. The model becomes fully Bayesian with the assignment

of prior probability distributions to parameters, reflecting the belief in their value before introducing new data. There are a number of methods available to formulate priors based on stakeholders' opinions (Spiegelhalter, Abrams et al. 2004). Beta distributions for instance are best suited for use as priors in parameters that represent a probability (e.g. test accuracies). Priors for these can be constructed by consulting expert opinion on their assessment of the mode and a threshold above which they are 95% sure that the chosen value is the true one. This information can be used to construct the relevant beta distribution using appropriate software[10]. The same software can be used for constructing beta distributions based on trial data where the number of successes and failures initially used to obtain the probability are available.

The priors assigned to the test sensitivities and specificities were based on previous evidence (Kolaczinski, Mohammed et al. 2004; Malik, Khan et al. 2004; Bell, Wilson et al. 2005; Reyburn, Mbakilwa et al. 2007) and expert opinion on the likely values for these parameters, and were expressed using beta distributions. They suggest that RDT1 has a lower sensitivity and higher specificity than RDT2. These estimates were then updated with the use of the trial data. The model was later run using a range of hypothetical priors to assess how these influence results.

The model categorises patients into four groups – True Positives, True Negatives, False Positives, and False Negatives, in each of the trial sites and for each RDT and presumptive treatment. The number of patients in, for instance, a true positive group is a function of a binomial distribution $TP \sim dbin(sens,par)$, where par is the total number of parasitaemic patients and $sens$ is the test sensitivity. Similarly for true negatives, $TN \sim dbin(spec,npar)$ reflects the likelihood for values of test specificity given the data on non-parasitaemic patients. The priors for these parameters were then updated by the data to provide posterior estimates for test sensitivity and specificity (Figure 6-3).

---

[10] 'BetaBuster' is available from The Graduate Group in Epidemiology, The University of California, Davis, http://www.epi.ucdavis.edu/diagnostictests/betabuster.html   Accessed February 7th 2008
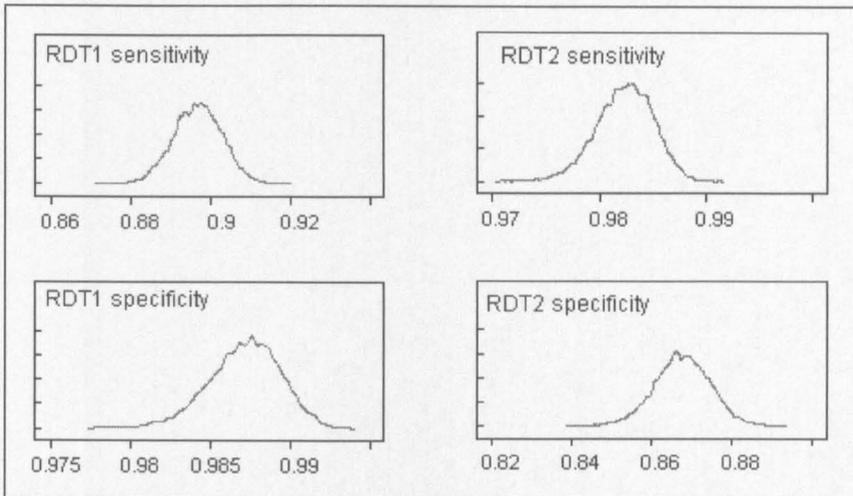
Figure 6-3 Posterior distributions for RDT sensitivities and specificities

The model assigns a cost to each of the four categories as in the deterministic model, using a gamma distribution instead of point estimates. The use of gamma distributions to represent cost data is considered more appropriate than normal distributions given the tendency of cost data to be highly skewed to the right (Fryback, Stout et al. 2001). Figure 6-4 is a simplified illustration of the model structure and parameters for all costs that follow the use of each RDT.

In contrast to the deterministic model and given current software limitations, there is no easy way for users to switch between age-groups without manually re-adjusting a number of parameter estimates, or duplicating the model with data series for all age-groups. Many of the input parameters would vary depending on the age of the patients; in this instance data were used that were relevant to adult patients.
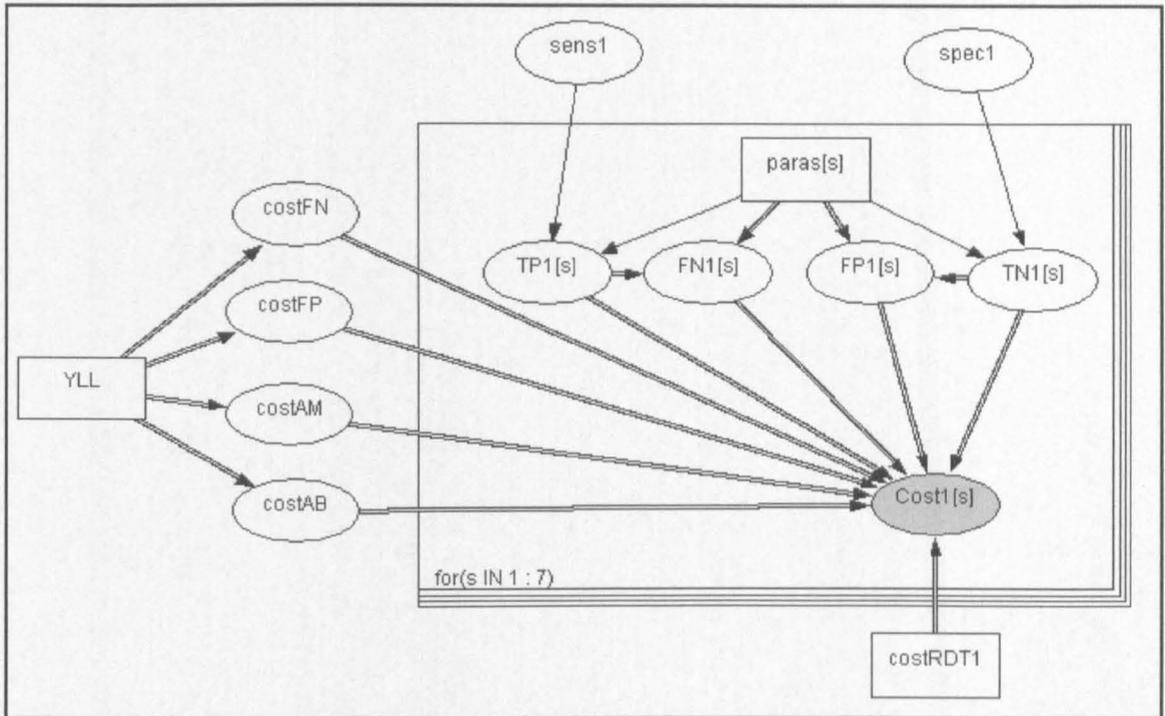
Figure 6-4: Model structure for estimation of total costs for each RDT. The parameters inside the plate are those that vary by site, notably the number of parasitaemic patients, and the number in each category (true positives etc.). Above the plate are the priors for test accuracy. To the left are other parameters that feed into the total RDT cost, notably 'costFN' and 'costFP' for the costs of false negatives and false positives respectively, and the value of life years lost on the far left that informs both the cost of patients wrongly diagnosed, and the total cost for antimalarials (costAM) due to adverse outcomes of treatment. 'par' is the total number of parasitaemic patients in each site.

Using a Markov Chain Monte Carlo simulation, the cost difference between RDT1, RDT2 and presumptive treatment is then repeatedly calculated, and by observing results over a high number of iterations, the probability that each of these strategies is most efficient is derived. Figure 6-5 depicts results for three locations with different transmission intensities, indicating that the probability that RDT1 is more efficient is over 80% in the low transmission site and diminishes with transmission intensity, whereas RDT 2 is more likely to be the most efficient option in the high transmission site. Presumptive treatment shows a probability of under 35% of being the preferred strategy in the high transmission site, and therefore is least likely to be the preferred strategy in this setting.
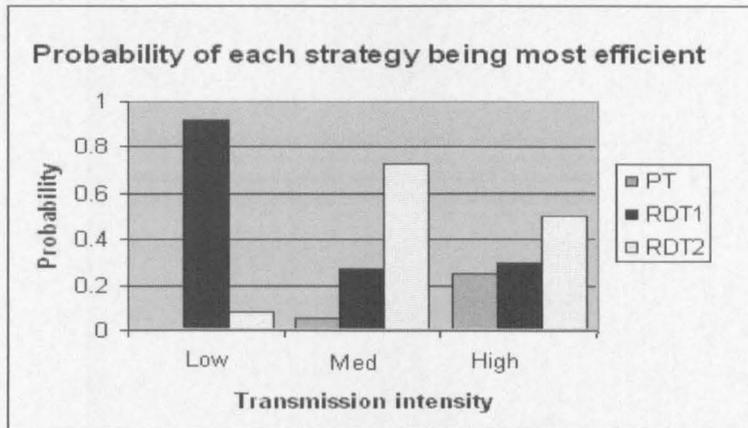
Figure 6-5 Results from the Bayesian model in WinBUGS (graph generated in Excel). The graph indicates the probability that either of the two RDTs or presumptive treatment is most efficient in three of the trial sites, with increasing transmission intensity in each.

One contentious issue that haunts Bayesian analyses is how to assign priors and the influence these exert on the data (Spiegelhalter, Abrams and Myles 2004). In this model, as is often the case, the influence priors had was relatively small, and only became apparent when setting these as extremely strong and divergent from the data. For instance only when the priors for the sensitivity and specificity of the RDTs were set at below 50% (i.e. flipping a coin would be a better predictor of the true status of the patient) with a very tight distribution around this estimate, did the use of presumptive treatment become more attractive at low transmission settings.

Discussion

Where cost-effectiveness is very context specific, as may be the case with malaria related interventions, decision support tools can help identify the most efficient use of scarce resources (van Gool, Gallego et al. 2007, English and Scott 2008). The use of decision support tools allows for the incorporation of local data and stakeholders' parameter estimates into the evaluation of intervention efficiency. This may enhance understanding of how variation in these influences results, and promote the intervention's adoption only where its benefits truly outweigh its costs. This method of evaluation differs considerably from the use of generalised evaluations, in that rather than the analyst aiming to provide

135

definitive generalisable results, policy makers are instead given a mechanism with which they can engage to explore the different options under consideration.

Technically, DSTs are similar to models used in standard evaluations, such as that in Chapter 5. The difference lies mostly in the interface of DSTs, which facilitates the incorporation of different parameter estimates, and changes to the model structure (e.g. changing perspective or the exclusion of factors like the harm of treatment should stakeholders choose to ignore this). More importantly however, the use of DST changes the focus of decision modelling from the provision of definitive results, to the development of adaptable models tailored to local circumstances. This places many decisions that were previously at the analyst's discretion, back into the hands of stakeholders and decision makers.

Previous economic evaluations of RDTs are few in number as shown in Chapter 2, and all but one (Shillcutt et al. 2008) drew on a single trial to produce their results, under the assumption that these would be informative in other settings. The Shillcutt et al. (2008) analysis was the first to draw on a range of secondary sources and the model was made available on the WHO website for users to apply to their own settings.

This chapter has presented two DST approaches to the incorporation and interpretation of local data in models, and the handling of uncertainty surrounding these. Both approaches can provide up-to-date and locally relevant decision recommendations but there are significant differences in the respective model structures and outputs.

The Bayesian model has higher conceptual and technical requirements. For this reason it may be considered less appropriate for practical use by decision makers. Its main advantage is that where there is considerable variation in parameter estimates between local data and previous evidence, the use of priors can moderate these differences according to the strength of each component. Where strong priors exist from other settings and only weak data exist for the location of interest, the priors will exert greater influence on the posterior distribution. For example, previous studies have found substantial variations in estimates of RDT accuracy (Cruciani, Nardi et al. 2004; Hopkins,

Kambale et al. 2007; Reyburn, Mbakilwa et al. 2007), so a single local study might still appear to be an insufficient basis for determining point estimates for parameter values. Use of the Bayesian approach will result in more conservative parameter estimates as the local data will be anchored by prior evidence from other sites. This might increase policy makers' confidence in results by ensuring that local data and parameter estimates are not entirely detached from a broader body of evidence.

As well as incorporating prior evidence into the analysis, the Bayesian model integrates all uncertainty simultaneously. This might be more appealing to policy makers interested in the 'bottom line' and not in the interaction of individual parameters and results. The model output can also provide policy makers with a degree of certainty in its predictions, allowing for the incorporation of risk-aversion in policy making. This functionality could also be built into the deterministic model, although at the risk of becoming overly complex for users who are less familiar with concepts such as probability distributions.

This trade-off between model complexity and the practicality of use is inherent to all decision models, but particularly so in the development of DSTs where it is a priority that decision makers can interact competently with the model. The main issue relates to potential disparities between the analytical, technical and conceptual demands when using the models, and the capacity of decision makers. As both the theoretical foundations and technical functions of decision models evolve, there will be ever greater challenges for decision makers engaging with 'black box' models and analyses. Keeping the model as simple and transparent as possible is likely to increase the confidence that decision makers have in the models. In addition to the greater conceptual demands of the Bayesian models, the technicalities of using software such as WinBUGS makes it less intuitive, so making use of such models would require ongoing expert support to users.

The deterministic model, on the other hand, appears more transparent as the manner in which variation and uncertainty in individual parameters influences results is easily visible. Experience of running the model with policy makers and stakeholders suggests that the

model facilitates their engagement in the estimation of input parameters[11]. The risk here is that extreme values might be selected, which may be considered accurate locally but have no grounding or justification from other settings. This contrasts with the use of probability distributions where more extreme values are less likely to be used, and the use of priors that draws the values to the confines of existing evidence.

The extent of the trade-off between model complexity and practicality is likely to diminish should decision makers become better acquainted with the concepts and methods used in economic evaluations, and also with developments in the software available to run DSTs. There are a number of programmes that complement Excel enabling the assignment of probability distributions to parameter estimates with 'user friendly' control of these. The integration with priors as carried out in WinBUGS however is not yet possible.

In addition to the limitations associated with each particular approach, there are a number of broader concerns regarding the ability of policy makers to make use of DSTs, and their acceptance of results in the face of other considerations. An assumption has been made regarding policy makers' familiarity with concepts such as the handling of uncertainty and even the use of probabilities, which might in fact not be the case. Even a basic understanding of economic evaluation may be absent, and economic evaluation is often confused with more limited cost analyses (Teerawattananon 2007). In this respect the use of DSTs can appear to be more of an ideal than a practical tool. Conversely policy makers might require more elaborate models that consider factors such as equity or disease severity in addition to efficiency. The main shortcoming of the work presented in this chapter is that it was not possible to formally test either or both of the models with policy makers and evaluate how they are received and used. This is a matter for further research.

---

[11] The model was presented at a Malaria Consortium workshop in Kampala where strategies for malaria diagnostics in Uganda were being evaluated, and later presented at the 2007 American Society of Tropical Medicine and Hygiene annual conference.

6.7 *Chapter conclusion*

Economic evaluations can provide valuable information to policy makers considering the adoption of new interventions. In a rapidly evolving and variable environment such as that of malaria transmission and control, it is unlikely that standard evaluations can maintain their relevance across different regions and over extensive periods of time. The use of DSTs allows for the incorporation of local and recent data, and stakeholders' parameter estimates, into the evaluation of intervention efficiency. This should enhance understanding of how variation in these parameter estimates influence results, and promote the intervention's adoption only where its benefits truly outweigh its costs.

The deterministic model facilitated the exploration of a variety of scenarios for different policy options and parameter estimates, offering a potentially valuable learning tool when contemplating policy options. The model is likely to be accessible to stakeholders who lack modelling expertise or support.

The Bayesian model had two methodological advantages: the incorporation of prior evidence and beliefs into the analysis, and the expression of uncertainty using probability distributions that carry through to model outcome. These are likely to contribute to best possible decision-making in practice. Conceptually and technically however, this approach is more demanding. Where DSTs are developed, the choice of model structure will have to balance methodological validity against the practical requirements and limitations of the decision making context.

All decision making practices have their shortcomings, and some of those relating to DSTs were described above. While these limitations are genuine, the use of even simple DSTs may represent a significant improvement over competing decision making practices, such as purely intuitive ones or reliance on outdated and less relevant evaluations. It is against these practices that DSTs should be evaluated. Future DSTs could be developed to consider additional factors beyond efficiency, such as equity weights, as an integral part of the analysis (James, Carrin et al. 2005). In this respect the DSTs presented here are relatively simple examples of an alternative form of economic evaluation, one that could be comprehensive yet versatile and adaptable to local and evolving circumstances.

7. Multilevel modelling for the evaluation of artesunate for the treatment of severe malaria

Multi-centre trials offer the benefit of greater generalisability of their results and are now routinely employed in assessing antimalarial effectiveness and cost-effectiveness (Adjuik et al. 2002; Dondorp et al. 2005; Chanda et al. 2007). They do, however, pose analytical challenges in how any variation in results between centres is accounted for. Multilevel modelling (MLM) can potentially address these challenges, but despite its methodological strengths it has been used in only a handful of instances in economic evaluations in general, and never in the context of malaria. The aim of this chapter is to explore the application of this approach to the evaluation of a multi-centre trial of antimalarials. The chapter provides a brief background to the use of econometric methods in economic evaluations in general, with a focus on the use of MLM in the context of multi-centre trials. In the subsequent section, MLM is applied to a trial comparing treatments for severe malaria. The data are analysed using standard methods and a series of regression models increasing in complexity building up to a random slope multilevel model.

7.1 Introduction to multilevel modelling

Randomized controlled trials (RCTs) are increasingly undertaken in a large number of sites, often located in different countries. In addition to the potentially larger sample size, such multi-centre trials aim to increase the degree to which their results can claim to be generalisable beyond any one particular trial setting (Grieve, Nixon et al. 2005). An increasing number of RCTs are also incorporating an economic component so they can assess not only the efficacy of the interventions, but also their costs and cost-effectiveness (Manca, Rice et al. 2005). The advantage of doing so is that this provides cost data that can claim greater internal validity as compared with gathering cost data either retrospectively or from secondary sources.

Consequently, a multitude of patient level data on cost and effectiveness from a range of settings becomes available. While this might appear to provide a greater degree of certainty in parameter estimates, there are a number of challenges to the inference of parameter estimates drawn from multi-centre data.

The simplest method to handle these data is to pool them into a single measure of central tendency with an expression of the uncertainty surrounding it. This takes full advantage of the entirety of observations, minimising the error term and confidence interval around the estimate. Having data from a range of sites can also suggest greater external validity as the inferences are made from a larger number of locations, suggesting the results should be a better representation of the true parameter value.

Using as an example a hypothetical multi-centre trial of an intermittent preventive treatment (IPT) intervention, the relationship between the number of doses given and the subsequent net monetary benefit can be explored. As Figure 7-1 shows, there appears to be a strong positive correlation between the number of doses given per year and the subsequent net monetary benefit. With a narrow standard error of 0.035 around the coefficient for the slope, the estimate appears very precise (95% CI 0.731-0.735).
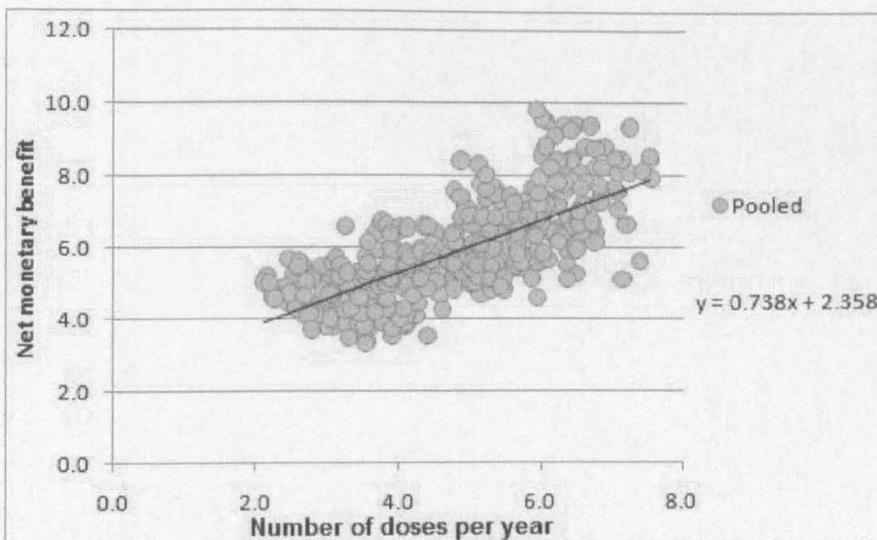
Figure 7-1 - OLS regression for the effect of treatment on Net Monetary Benefit

This pooling, however, can mask genuine variation between the centres where the trial took place (heterogeneity). Transmission intensity for instance is likely to have a significant impact on individuals' initial susceptibility to infection, and how likely they are to develop severe illness (Marsh and Snow 1999). Similarly, provision of IPT might be more costly in areas where the population is harder to access. Furthermore, it might be found that there is an interaction between these factors, so that areas that are harder and more costly to access systematically feature higher prevalence of malaria, as is known to occur when comparing rural and urban settings in similar geographical regions (Roca-Feltrer, Carneiro et al. 2008). Pooling these results could lead to incorrect inferences, and unduly narrow error terms around these estimates (Rice and Jones 1997; Grieve, Nixon et al. 2005).

Pooling data ignores any hierarchical structures that might be present, as individuals for instance can be clustered within centres, and centres within countries, which for any one of a variety of reasons may influence outcomes of interest. Costs, for example, either for existing or new interventions, can vary across centres and countries (Grieve, Nixon et al. 2005). Hospital financing structures can influence the case mix, for instance. Where patients are charged per service, the case mix is likely to be more severe, influencing the intervention effectiveness.

Returning to the IPT example, Figure 7-2 depicts the same hypothetical data shown in Figure 7-1, in this instance broken down by centres. As is evident, differences between centres can be seen not only in the point estimates for particular coefficients, but also in the direction of the association and in the variance of individuals around the centre mean, implying different strengths of association.
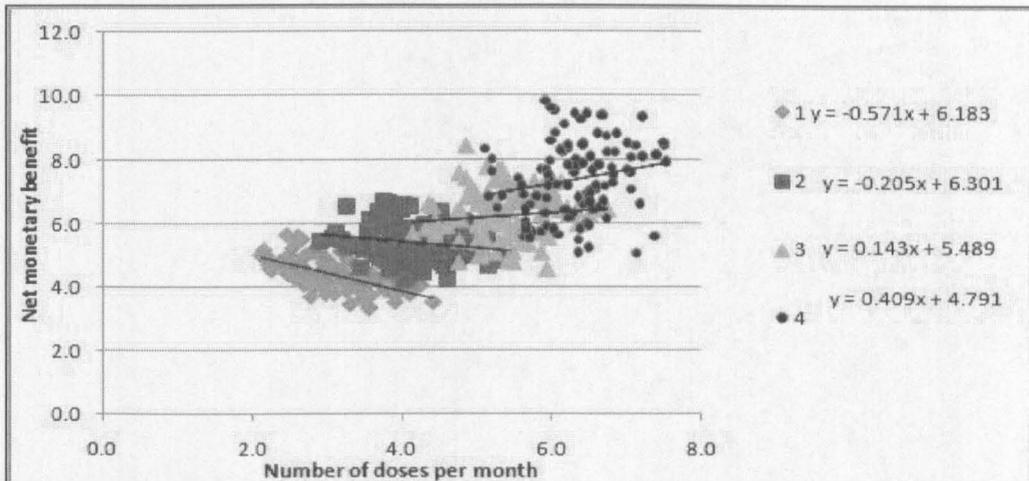
Figure 7-2: The effect of treatment on net benefit in each site varies considerably, as indicated by the regression lines for each of the centres. In centre one there is a strong negative association between the number of doses and the net monetary benefit. Centre four on the hand has a higher net benefit with a positive correlation, although the individual observations appear more widely dispersed

The figure suggests that the pooling of the data might lead to erroneous correlations that would be further supported by the narrow error term attached. In Figure 7-2 on the other hand, the data are completely stratified. This stratification, however, implies that a smaller number of observations are available for inference. In contrast to the pooled estimate that was supported by narrow error terms, here the error terms are wide, as seen in Table 7-1.

		Coefficient	SE	Lower 95%	Upper 95%
Centre 1	Intercept	6.18	0.24	5.70	6.67
	Slope	-0.57	0.08	-0.72	-0.42
Centre 2	Intercept	6.30	0.43	5.44	7.16
	Slope	-0.21	0.10	-0.41	0.02
Centre 3	Intercept	5.49	0.80	3.91	7.07
	Slope	0.14	0.15	-0.15	0.44
Centre 4	Intercept	4.80	1.30	2.65	7.81
	Slope	0.40	0.21	-0.07	0.74

Table 7-1: Regression coefficients and the uncertainty surrounding these for each centre individually
The coefficient for the slope in this example represents the incremental net benefit, therefore where this is negative the intervention is not considered efficient. The loss of power following stratification means that only the first of the four sites has a coefficient for the slope that is statistically significant at the $p=0.05$ level, and the analysis as a whole provides ambiguous results. Furthermore, where differences between sites are evident, an

143

analysis that stops short of explaining these can appear incomplete. Most importantly, stratifying the results implies that the aim of attaining generalisable results may be lost.

This illustrative example is perhaps an extreme one, where the association between the explanatory and dependent variable changes from positive and strong, to an uncertain one, once the data are stratified by centre. The issues raised however are likely to underlie most multi-centre analyses, where a choice has to be made between either pooling data and masking differences between sites, or stratifying results and leaving the reader to decipher how significant any differences might be between them, their causes, and how to interpret the greater uncertainty attached to the estimates.

Multilevel Modelling offers a third approach, drawing on the individual observations available across all second level units, without making an assumption that these are all independent observations by recognizing their hierarchical nature (Manca, Rice et al. 2005) The advantages of using MLM can be summarized as follows[12]:

1. When obtaining pooled estimates, the error around these will be broader, reflecting the heterogeneity of the parameters at the different sites

2. For stratified results, use of MLM will result in both a certain convergence of the site coefficients towards the mean, and also in shrunken error terms around the estimates as these will borrow strength from the other sites

3. Covariates can be used to explain the variation in different sites; these can then be used to predict results in other settings. As opposed to single level multiple regression, the use of MLM ensures that these explanatory variables are associated with the appropriate level

---

[12] Much of the information in this chapter on MLM was gathered from an extensive online course made available by the University of Bristol Centre for Multilevel Modelling
http://www.cmm.bristol.ac.uk/learning-training/course.shtml   Accessed July 2008

At the heart of MLM lie the relationships between three different types of residuals that reflect the dispersion of the data within and amongst the higher level units. These can be illustrated graphically as the vertical differences between a) the individual observations and the pooled regression line (the left-most arrow in Figure 7-3); b) the average distance between the centre regression line and the pooled one (middle arrow); c) the distance between each data point and its respective centre regression line (right most arrow). Each of these sets of residuals is summed up in a distribution, the mean of which is by definition zero (as the explanatory variables seek to determine the value of the dependent variable while the residual is the unexplained variance that cancels each other out).
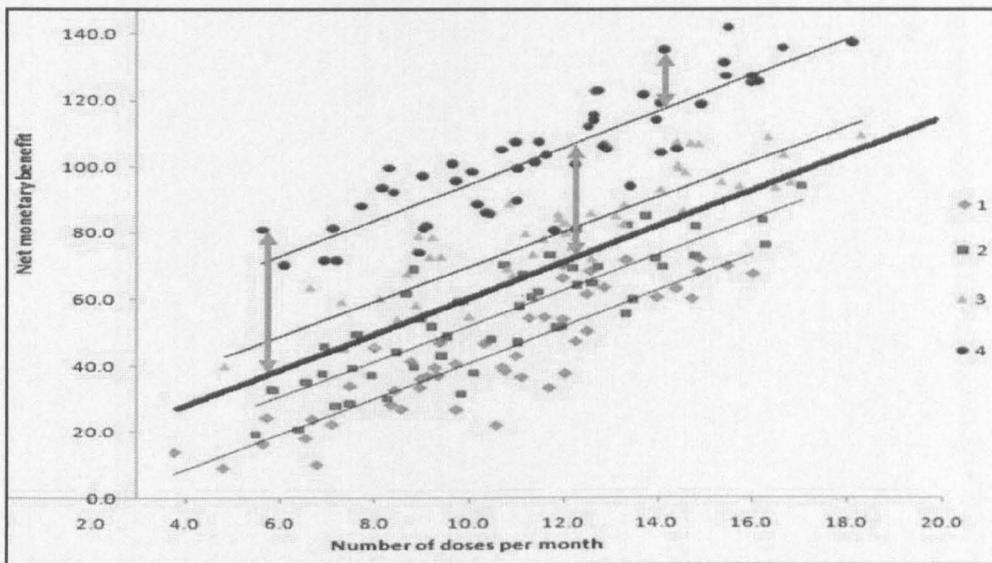
Figure 7-3: Pooled and stratified regression lines; the arrows depict the different residuals. This dataset depicts a simpler scenario than the previous one. Here, the incremental net benefit of IPT is assumed to be fixed, as indicated by the parallel slopes, although the baseline for each site is different as indicated by the vertical differences between the site average slopes.

One of the advantages of MLM is the ability to infer how much of the variation, and therefore the association, is derived from the clustering of the data around the individual site mean, as opposed to their association with the predicted pooled effect. This relationship is summarised in the intra-class coefficient (ICC), a product of the centre variation divided by the total variation, as shown in equation 1, where σ_μ^2 is the centre variation around the overall mean and σ_ε^2 is the variation of the individual observations around the overall mean

145

$$ICC = \frac{\sigma_\mu^2}{\sigma_\mu^2 + \sigma_z^2} \qquad (1)$$

Most importantly, with a better understanding of the strength of these different associations, MLM can infer the degree to which the estimates of the coefficients for each site should be influenced by the data from all other sites. Where for instance the number of observations in a particular centre is low, or the variance around the mean is high, the estimate for the site would be pulled in to a large extent towards the pooled estimate. This is termed *shrinkage*, referring to both the convergence of site effects towards the pooled effect (Figure 7-4), and also to the reduced measures of uncertainty surrounding the individual centre estimates that borrow strength from the pooled data.

Conversely, the pooled estimate increases its confidence intervals, as the model now recognizes the heterogeneity in site effects and the inappropriateness of pooling the data without accounting for this variation. The degree of both the shrinkage in individual site variation and the increased uncertainty of the pooled estimate is a product of the different components of the ICC.
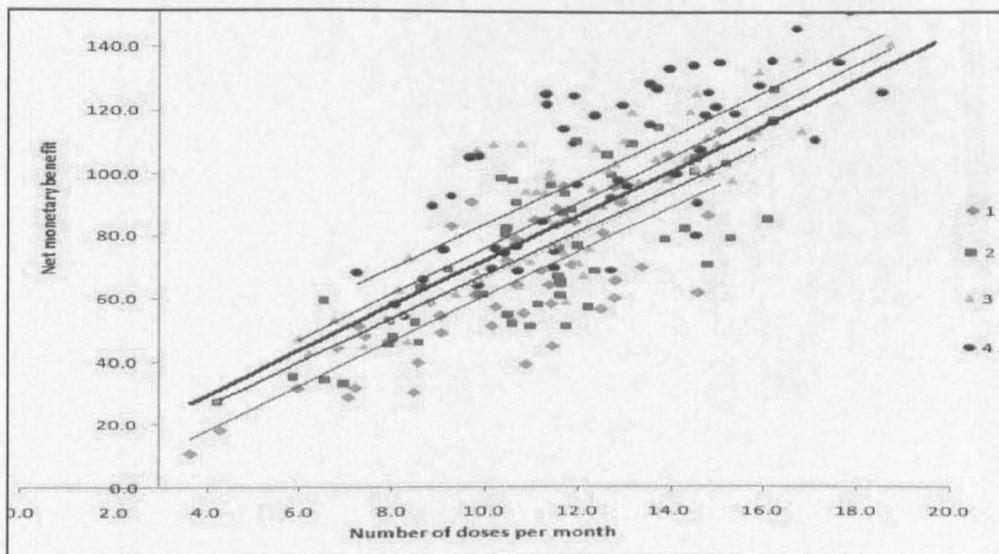
Figure 7-4: The individual centre regression lines are less dispersed than in the stratified results in Figure 7-3

146

7.2 How does MLM differ from multiple regression analysis?

To a certain extent hierarchical structures can be modelled with single level, multiple regression, using either the higher level units themselves as explanatory variables, or factors that are generally determined at that level. If, for instance, it is thought that treatment outcome is likely to be influenced by factors relevant at the higher level, such as HIV prevalence in each country, this can be addressed by either including a variable that represents the country effect on outcome, or by introducing a variable representing HIV prevalence for each individual observation.

Introducing dummy variables for each country will estimate the effect these have on outcome (a *fixed effect* model) by controlling for this in estimating the treatment coefficient. This is essentially an analysis of variance (ANOVA). There are a number of limitations to this approach. First, if the number of countries in the analysis is high, the model becomes unwieldy, particularly if the analysis aims to estimate not only the country influence, but also the interaction with other factors (e.g. level of health care facility), in which case the large number of coefficients is unlikely to produce robust results. Second, it can be shown that it is impossible to simultaneously control for higher level clusters (countries; hospitals etc.) and at the same time include factors that are determined at that level in the model, such as HIV prevalence (Rice and Jones 1997).

While the use of a series of dummy variables to represent for instance a range of countries is essentially multiple regression, these variables all represent different values for a single categorical variable. One of the main advantages of using regression over standard methods is that a range of different factors can be included in the analysis, both to explain some of the remaining variance after treatment is accounted for, and also to control for other factors that might for instance be confounders in the apparent treatment effect. The extension of the pooled model is simply a matter of adding the variables of interest and estimating their coefficients. The model's goodness of fit can then be tested as compared with the simpler model.

There are however limitations on the number of variables that can be included if the model is to maintain its validity. First, variables to be included should be stated in advance, as apparently convincing correlations may arise by chance; the inclusion of large numbers of variables in the hope of identifying correlations is a likely cause of type I errors. Second, the inclusion of additional variables weakens the strength of the association for existing ones, as the number of observations relative to the number of variables diminishes (Kirkwood and Sterne 2003). These restrictions are true regardless of hierarchical data structures. Where such structures are known to exist, however, there is an additional concern around the use of single level multiple regression. This relates to the inclusion of factors that are associated with higher level units being assigned to lower level ones, and giving them undue strength in a single level analysis.

Most importantly however, is the assumption underlying multiple regression (and all single level models), that the observations are independent of each other – in other words that they are not clustered in higher level units. As was shown with the IPT example, where this is not the case the use of single level model can result in erroneous inferences.

In the remainder of this chapter the use of MLM will be compared to a standard analysis of cost data for a multi-centre trial for the treatment of severe malaria.

7.3 Use of MLM in estimating the cost of treatments for severe malaria

7.3.1 Background

ACTs are now recommended for first line treatment of uncomplicated falciparum malaria in all malaria endemic countries. They have repeatedly been shown to be more effective and cost-effective than their predecessors (Honrado, Fungladda et al. 1999; Agnamey, Brasseur et al. 2005; Wiseman, Kim et al. 2006; Yeung 2006), require only once a day dosing and are associated with few adverse effects (Johann-Liang and Albrecht 2003). For severe malaria, quinine has been the traditional "gold-standard" treatment in both

developed and developing countries (Dondorp Nosten et al. 2005). Quinine is effective, but it is not simple to administer and it has a narrow therapeutic ratio. It is associated with a significant risk of local toxicity following intramuscular injection, and significant risks of systemic toxicity (hypoglycaemia, hypotension if administered rapidly). Quinine must be given three times daily either by constant rate intravenous infusion or intramuscular injection to the anterior thigh, a painful and potentially damaging procedure (Anstey, Price et al. 2006).

A growing body of evidence, summarised in recent reviews, demonstrates the considerable superiority of artesunate relative to quinine in terms of mortality rates without an increase in rates of adverse outcomes (Cochrane review estimate; RR 0.62, 95% CI 0.51 to 0.75 (Jones, Donegan et al. 2007)). The studies so far have included mostly adults in Asia, although of the 1461 patients enrolled into the large multi-centre SEAQUAMAT trial, 202 were children, for whom results were similar (Dondorp, Nosten et al. 2005).

The SEAQUAMAT study was conducted across ten sites in four South East Asian countries. In total, mortality in patients treated with artesunate was 35% lower than in quinine recipients. The implication was that for every 13 patients treated with artesunate instead of quinine, one death would be averted. Despite these promising results, and endorsement by the WHO treatment guidelines (WHO 2006), even within Asia most local guidelines continued to specify quinine as the drug of choice for severe malaria. The second most frequently recommended treatment for severe malaria is artemether, even though its advantage over quinine in terms of mortality has been shown to be limited (Hien, Day et al. 1996; Pittler and Ernst 1999). Artesunate has only recently been added to the policy guidelines of a limited number of countries in Asia (WHO 2007).

In order to explore the cost-effectiveness of artesunate for the treatment of severe malaria, an economic evaluation was carried out using the SEAQUAMAT trial data as a foundation for the analysis. This was carried out using standard methods, averaging costs and health outcomes to calculate the cost per death averted for the use of artesunate. The evaluation concluded that artesunate is a highly cost-effective intervention, as the cost of averting a death was below $150 (see Annex 3).

The evaluation, however, did not account for the variation between sites, and this was identified as an area for further research using MLM. Ideally the measure of outcome for such an analysis would be the net-benefit as this encompasses both cost and health outcomes. The distribution of net-benefit in this instance however is bimodal, and does not approximate a Gaussian one, and so is not amenable to regression analysis.

The aim of this analysis was therefore to examine the costs of switching from quinine to artesunate from an economic perspective. This is explored by comparing the results for a pooled analysis, with a stratified one, and finally using a multilevel model. Differences in costs can be of high relevance for policy makers who might be deterred from switching to artesunate without a clear idea of its financial implications.

7.4 *Methodology*

Interventions. The interventions being considered were quinine and artesunate for treatment of severe malaria. The drugs were given intravenously.

Trial data. The SEAQUAMAT study was carried out between 2003-2005 in one site each in Bangladesh, India, Indonesia and seven sites in Myanmar (Dondorp, Nosten et al. 2005). Relevant data from the trial for this analysis were drugs and dosages used, the equipment needed to administer the treatments, and the length of stay in hospital as inpatients (WHO 2008).

Cost data. Costs for artesunate were obtained from the producer and included shipment costs. Quinine costs and those for i.v. sets and syringes to administer the drugs were obtained from the International Drug Price Indicator Guide (MSH 2007). Drug costs were increased by 15% to account for taxes and an extra 10% for wastage (Gold, Siegel et al. 1996). Standard inpatient care costs for each country (excluding drugs) were obtained from the WHO-CHOICE database; these included "hotel" costs – those for personnel, capital and nourishment [22]. These costs were then assigned to each individual patient, according to the dosage they were given and their length of stay as an inpatient.

It was assumed that apart from the cost of trial drugs, the inpatient cost per day was the same for both treatment arms in each site. Labour costs were also assumed to be equal, although artesunate is simpler to administer (Anstey et al. 2006). Costs were converted from local units to US dollars at the relevant year, adjusted for inflation using the consumer price index, and reported in 2008USD. Table 7-2 shows the values used.

Item	Unit cost		Source and notes
Quinine vial	$0.19		International Drug Price
Quinine tab	$0.04		Indicator Guide. Accessed 2/6/08
Artesunate vial	$1.2		Quote from the producer
Artesunate tab	$0.17		
Cost per inpatient day	Bangladesh	$4.6	WHO-Choice estimates by
	India	$8.4	country and hospital level.
	Indonesia	$2.0	Accessed 23/5/08
	Myanmar (mean of all Myanmar sites)	$1.7	
Quinine administration equipment	$1.2		1x 5ml syringe and 2 x needles, 1 x infusion set, 1 x (500ml or 1000ml) bag of IV solution
Artesunate administration equipment	$0.3		1x 5 ml syringe, 2 x needles,

Table 7-2: Costs for treatment, equipment and inpatient care used in the analysis

Analysis

The framework was a cost analysis used to determine the incremental cost per inpatient with severe malaria treated with artesunate instead of quinine. Differences in costs between the two arms were initially summarised by averaging all costs, both pooled and stratified by site, and then re-calculated using regression analysis and MLM for comparison.

Regression models. The first model used was an ordinary least squares regression for costs dependent on treatment, pooling patients from all sites (Equation 2). These results should approximate those used in standard calculations of differences in pooled costs:

$$y_i = \beta_0 + \beta_1 x_i + \varepsilon_i \quad (2)$$

With this specification, x is a dummy variable representing the treatment arm for patient i, the intercept coefficient (β_0) represents the cost for the baseline arm, in this instance the use of quinine, while the coefficient for the slope (β_1) provides the incremental cost for the use of artesunate. Use of the regression framework provides an elegant summary of the uncertainty surrounding the values, which are a product of the variation between the individual data points and the regression line and the number of observations. This uncertainty is summarised in the standard errors for the two coefficients and can be expanded with the use of chosen confidence intervals to assess whether the correlation is a statistically significant one. This uncertainty differs from the error term (ε_i) which is the residual distance for each individual patient from the regression line.

For reasons discussed above, pooling the data when these originate from a hierarchical structure can lead to imprecise inferences. The first possible adaptation, still using a single level model, is to stratify the observations by introducing a range of dummy variables that represent each of the sites. These variables are assumed to control for factors at the site level, without explicitly stating what these are in the model. The result is an extended model where each of the sites (apart from one chosen reference site) is represented by a variable that takes on a value of 0 or 1 (Equation 3).

$$y_i = \beta_0 + \beta_1 x_i + \beta_2 x_2 + \beta_3 x_3 \ldots \beta_n x_n + \varepsilon \quad (3)$$

Use of the stratified model becomes restrictive when introducing other factors, firstly since they become unwieldy, and more so due to possible interaction between the other factors and the effect of the sites. These problems are avoided with the use of MLM, where the model structure reflects the hierarchical structure of the data, and where higher level units are assumed to be random samples of a broader population. The first MLM was a random intercept one, which assumed that the base cost per patient treated with quinine varies by

152

site, while the use of artesunate incurs the same additional costs in all sites. In Equation (4) the subscript of y is now ij, where i represents an individual patient in centre j. The fixed part of the model consists of the same fixed elements as Equation (2), while the additional random term u has been introduced representing the random variation in the intercept by site.

$$y_{ij} = \beta_0 + \beta_1 x_i + \varepsilon_i + \mu_j \qquad (4)$$

The assumption however that β_1 is fixed can be restrictive, therefore allowing it to vary can indicate whether its effect varies significantly by site and the extent to which the data can be aggregated. The random slope model produces an additional parameter which is a measure of the covariance between the slope and the intercept; this can determine whether for instance sites that have higher costs for quinine might experience even higher costs for artesunate, thus fanning out the regression lines.

Testing for model fit. A number of tests are available to examine the significance of individual coefficients and of the model in totality. One method uses the -2loglikelihood statistic, which is compared across the different models being assessed, and combines this with their degrees of freedom (based on the number of parameters in each model). These values are then tested using a chi squared test of significance (Centre for Multilevel Modelling 2008). The limitation of this approach is that models have to share structures in order to be compared. Furthermore, the calculation of the degrees of freedom is done without consideration of whether these factors do in fact have any effect in the model.

More recently the Deviance Information Criterion (DIC) has been developed for comparisons of models that differ in structures such as number of explanatory variables, distributional assumptions, and number of levels (Spiegelhalter, Best et al. 2002). The DIC functions by examining the deviance of the data from the predictions, and penalizes for model complexity in a single stage, rather than the 2 stage process required for the -2loglikelihood statistic. Furthermore, the DIC estimates the number of effective degrees of freedom, and penalizes the model only for these. Lower DICs indicate a better model fit. The DIC was compared for each model to identify the most appropriate one.

The regression analyses were carried out using MLwiN 2.10.

7.5 Results

Treatment costs – standard summary. The differences in treatment costs by site, the uncertainty surrounding the point estimates, and the pooled result are shown in Figure 7-5. The pooled estimate shows quinine to be less costly, with a 95% confidence interval that clears the vertical axis, suggesting a statistically significant result.   However, the figure also shows substantial differences between the sites, with four out of ten sites having a lower mean cost for artesunate than quinine. The uncertainty within each site also appears to be high, with seven of the sites having confidence intervals that cross over the vertical axis of zero difference. Inferring from the pooled estimate that artesunate is significantly costlier seems, therefore, questionable.
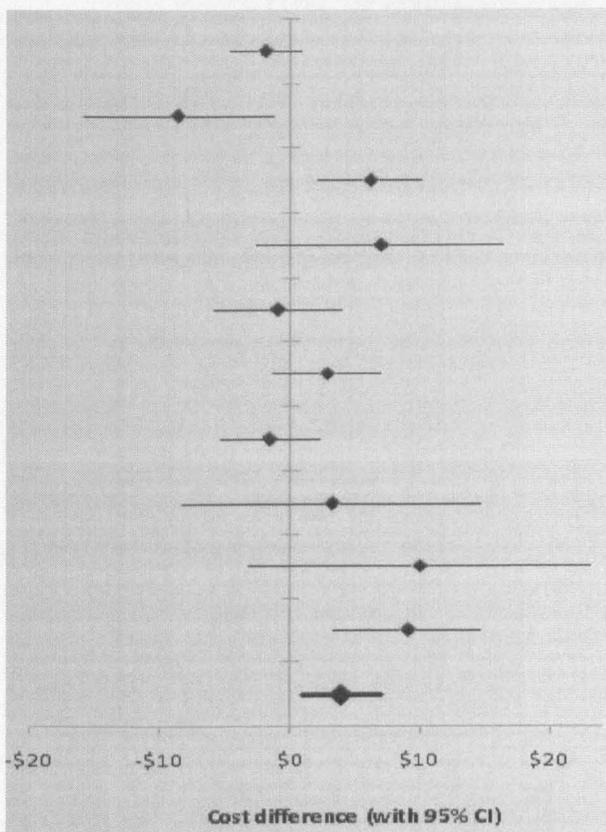
Figure 7-5: Differences in treatment costs and 95% confidence intervals across all sites and the pooled estimate (bottom).

154

7.5.1 Regression analyses

The simplest regression model pools the data from all sites and uses the treatment as an explanatory dummy variable and cost as the dependent variable. The output in MLwiN is shown in Table 7-3. In the regression equation, the intercept is the predicted cost for the use of quinine. The cost for artesunate is that for quinine plus the second coefficient, which represents the difference in cost between them. These results are almost identical to those obtained in the standard analysis. The standard error for the incremental cost of artesunate suggests that there is a statistically significant difference in costs between the two treatments.
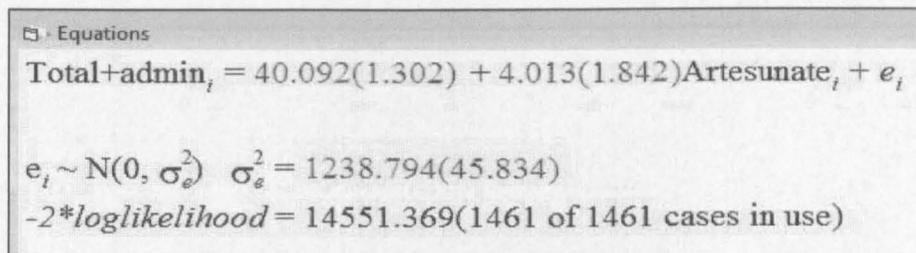
□ · Equations

$\text{Total+admin}_i = 40.092(1.302) + 4.013(1.842)\text{Artesunate}_i + e_i$

$e_i \sim \text{N}(0, \sigma_e^2) \quad \sigma_e^2 = 1238.794(45.834)$

$-2*loglikelihood = 14551.369(1461 \text{ of } 1461 \text{ cases in use})$

Table 7-3: MLwiN output for total cost regressed on treatment using a single level model for the pooled data. The top line shows the estimates for the coefficients, with their standard errors in brackets. e_i is the error terms for the individual observations, which is shown to assume a normal distribution (N) with variation of σ_e^2.

The random term represents the departure of each data point from the predicted regression line. In the second line this is shown to assume a normal distribution with mean 0 and variance of 1239, i.e. a standard deviation of 35.2. The bottom line is the -2*loglikelihood statistic which can be used to compare the model fit with other similar models.

The same model was used for each site individually, producing almost identical estimates to those shown in Figure 7-5 for the individual sites.

The second model used is an ANOVA, where the effect of each site is accounted for by introducing dummy variables to represent each one of them, apart from one reference site, in this case that in Bangladesh (Table 7-4). The model controls for the effect of sites on baseline differences, but assumes a fixed incremental difference for treatment with artesunate. Thus for the reference site, Bangladesh, the use of quinine costs $37.3 per patient, while artesunate implies an incremental cost of $4.2. For India, the incremental cost is significantly higher, adding a further $25.9. The standard errors for all coefficients indicate significant differences between almost all sites and the reference one, Bangladesh.

$$
\begin{aligned}
\text{Total+admin}_i = {} & 37.332(1.665) + 4.221(1.630)\text{Artesunate}_i + 25.869(2.990)\text{India}_i + 25.959(2.340)\text{Indonesia}_i + \\
& -1.184(3.436)\text{Myn.1}_i + -18.125(3.711)\text{Myn.2}_i + -16.561(4.442)\text{Myn.3}_i + -4.092(5.251)\text{Myn.4}_i + \\
& -23.643(4.440)\text{Myn.5}_i + -11.225(4.209)\text{Myn.6}_i + -14.190(2.718)\text{Myn.7}_i + e_i
\end{aligned}
$$

$e_i \sim N(0, \sigma_e^2) \quad \sigma_e^2 = 966.247(35.750)$
$-2*loglikelihood = 14188.206(1461 \text{ of } 1461 \text{ cases in use})$

Table 7-4: MLwiN output for a single level model with dummy variables representing each of the sites. The top line shows the estimates for the coefficients, with their standard errors in brackets.e_i is the error terms for the individual observations, which is shown to assume a normal distribution (N) with variation of σ_e^2.

This model and the previous ones were compared to see which better fits the data. The -2*loglikelihood statistic for the second model is indeed lower, therefore differentiating between sites has explained some of the variability in the data. The superiority of this model is confirmed by subjecting the difference between the two statistics to a chi squared test with 9 degrees of freedom (the number of variables in the model minus one). This produces a p value of <0.0001.

The next model fitted was a multilevel random intercept one. As shown in Table 7-5, this contrasts with the previous model where each site was represented by a dummy variable with its fixed effect estimated. In this instance the sites were assumed to be drawn from a larger population of sites, with a variance summarised in the added random term μ. The subscript 0 indicates that this relates to the randomness of the intercept (β_0), and

156

subscript *j* refers to each of the sites. In contrast to the previous ANOVA model, the MLM model provides an overall estimate for the incremental cost of artesunate, $4.2, and the 95% CI can be calculated as $1 to $7.4. These estimates are almost identical to those of the pooled model. Looking at the -2*loglikelihood statistic, however, this model does not appear to be a better fit than the ANOVA one.

$$
\text{Total+admin}_{ij} = \beta_{0j} + 4.221(1.635)\text{Artesunate}_{ij} + e_{ij}
$$
$$
\beta_{0j} = 33.830(5.329) + u_{0j}
$$

$$
u_{0j} \sim N(0, \sigma_{u0}^2) \quad \sigma_{u0}^2 = 265.802(125.059)
$$
$$
e_{ij} \sim N(0, \sigma_e^2) \quad \sigma_e^2 = 972.833(36.115)
$$
$$
\textit{-2*loglikelihood} = 14232.328(1461 \text{ of } 1461 \text{ cases in use})
$$

Table 7-5: MLwiN output for the random intercept model. The top line shows the estimates for the coefficients, with their standard errors in brackets. e_{ij} is the error term for individual observations within each centre *j*. u_{0j} represents the level 2 random departures (or residuals) of the sites from the mean intercept. σ_e^2 is the variance for the individual residuals and σ_{u0}^2 the variance for the site residuals

There are now two terms for the variances, one for the individual residuals around the site mean, σ_e^2, and one for the site residuals around the overall mean, σ_{u0}^2. The ratio $\frac{\sigma_\mu^2}{\sigma_\mu^2 + \sigma_e^2}$ provides the intraclass correlation, i.e. the percentage of the overall variation explained by differences between sites. In this instance ICC=0.21, therefore 21% of the variation is due to differences in cost between sites, rather than random variation of individuals.

Both the previous models allowed the baseline to vary (i.e. the cost for quinine), but assumed the incremental cost for the use of artesunate to be fixed. The final model to be fitted was a random slope model, where this restriction was relaxed and the coefficient for the slope was also supplemented by a random element.

$$\text{Total+admin}_{ij} = \beta_{0j} + \beta_{1j}\text{Artesunate}_{ij} + e_{ij}$$
$$\beta_{0j} = 34.329(5.047) + u_{0j}$$
$$\beta_{1j} = 3.187(1.920) + u_{1j}$$

$$\begin{bmatrix} u_{0j} \\ u_{1j} \end{bmatrix} \sim N(0, \ \Omega_u) \ : \ \Omega_u = \begin{bmatrix} 236.525(113.770) \\ 27.914(30.392) & 7.711(14.011) \end{bmatrix}$$

$$e_{ij} \sim N(0, \sigma_e^2) \quad \sigma_e^2 = 970.972(36.131)$$
$$-2*loglikelihood = 14230.942(1461 \text{ of } 1461 \text{ cases in use})$$

Table 7-6: MLwiN output for the random slopes model. The top line shows the estimates for the coefficients, with their standard errors in brackets. e_{ij} is the error term for individual observations within each centre j. u_{0j} represents the level 2 random departures (or residuals) of the sites from the mean intercept, while u_{1j} represents the random departure from the mean slope, and both terms are assumed to follow a normal distribution, with Ω_u estimating the covariance between the two residuals. σ_e^2 is the variance for the individual observation errors.

In this model, the incremental cost is slightly lower than the previous models, and most importantly the difference is no longer statistically significant (p=0.11). The random slope model output also shows that there is no significant covariation between the intercept and the slope, therefore increases in the cost of quinine were not correlated with even higher incremental cost for artesunate.

Figure 7-6 shows a comparison of all the model outputs. First, the pooled data shows that quinine is less costly than artesunate, as indicated by the downward sloping regression line. This result is identical to the standard method of averaging pooled data. The next panel shows the results of single level regression for each site individually, where the incremental costs appear to vary widely between the sites. The ANOVA method assumes a fixed baseline, but then provides a completely stratified incremental cost for each site, therefore the dispersion is wide. In the random intercept model, the sites borrow strength from each other according to the ICC, therefore they have shrunk towards the overall mean as in panel A. Finally the random slope model allows for different baseline costs, although the variance between sites in relation to the variance within sites is small as the regression lines are almost parallel. Having been moderated by data from other sites, overall the sites are less dispersed.

158
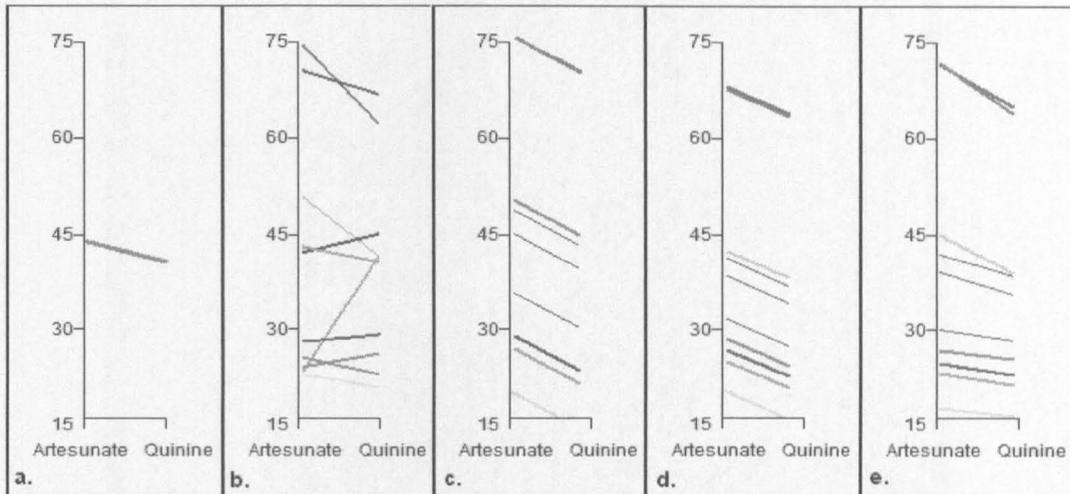
Figure 7-6:Graphs representing the incremental cost by: a. pooling the data; b. Stratifying the data; c. using ANOVA; d. a multilevel random intercept model; e. a random slope model

7.5.2 Testing model fit using DIC

The DIC for each of the models reviewed above are presented in Table 7-7, showing that the DIC reduced in size with the MLM models, with a very slight advantage for the random slope model. There are general rules of thumb as to what constitutes a significantly better model based on differences in DICs (Spiegelhalter and Best 2002). A difference of 1 or 2 implies that the models have a close to equivalently good fit. Differences of 3 to 5 are substantial enough to suggest that the model with the lower DIC is a superior fit.

The difference between the two MLM models, for instance, is small enough to consider both models in obtaining results, while they both supply considerably superior inferences to both the ANOVA and pooled models.

Model	Incremental cost	Degrees of freedom	DIC
Pooled	4.013	3	14557
ANOVA	Varies by site	10.82	14217
Random intercept	4.221	11.96	14213
Random slope	3.187	11.71	14211

Table 7-7: DIC for the models compared

159

7.6 Discussion

The aim of this chapter was to explore the use of MLM in the context of a multi-centre malaria trial. The analysis presented a series of models for the comparison of costs of treating severe malaria with either quinine or artesunate. The analysis demonstrated that use of MLM provided results that were different from those of a standard analysis, with implications for subsequent decision recommendations.

Whereas in a standard pooled analysis artesunate was found to be more expensive, with a high degree of certainty attached, the use of MLM showed this difference to be smaller and no longer statistically significant. In dozens of countries where average health expenditure is below $13 per capita per year (Jha and Mills 2002), an intervention that is expected to cost $4 more than its existing comparator may be viewed as prohibitively expensive. Policy makers who might have opted to reject the use of artesunate due to its higher costs as shown in a standard analysis, may find that a smaller difference, and one that is not statistically significant, may make the drug switch more acceptable.

MLM was also shown to influence the cost estimates of the individual sites, as these had shrunk towards the overall mean. This thesis has consistently argued that economic evaluations of malaria diagnostics and treatments must account for site variation in order to provide best decision recommendations. Where trials are carried out in a multitude of sites, policy makers might be interested more in the data arising from their own setting. Use of MLM allows for the incorporation of individual site results with those of other sites, while accounting for their individual variance and weight to determine the degree to which they should be drawn towards the overall mean.

MLM has not previously been used in the context of the evaluation of malaria related interventions, and has only been used in a handful of economic evaluations of health care interventions in general (Manca, Rice et al. 2005). Its use, however, has the potential to address many of the shortcomings of analyses that pool data clustered in higher level

units. The danger in pooling data is that it ignores the variability between sites and provides erroneously precise estimates of the measure of outcome. This is a particular problem in the context of malaria, where high variability can be present in both susceptibility to illness due to malaria prevalence and population immunity, and in the effectiveness of treatment due to factors like antimalarial resistance.

7.6.1 Limitations

In retrospect the dataset used for this analysis was not the most appropriate to explore the use of multilevel modelling. Firstly, one of the main advantages of using MLM, and indeed regression analysis as a whole, is the ability to introduce a range of factors that are assumed to influence cost and outcomes, both at the individual and at higher levels. MLM is particularly suited to represent these factors by placing them at the appropriate level. When this is not done, as in single level models, the effect of these factors is overestimated (Rasbash, Steele et al. 2005). In this analysis, however, no other factors were available that could be used as explanatory variables in the cost analysis, because the trial did not plan to collect cost related data, and most the inputs in the analysis were collected retrospectively.

A second limitation related to this is that the number of sites in the analysis was at the lower end of what is considered applicable to MLM. A larger number of higher level units would allow exploration of other factors that influence the difference in costs between the two treatments.

Third, ideally net-benefit would be used as the primary measure of outcome, capturing both costs and health consequences, and incorporating the decision threshold into the analysis. As the net-benefit in this instance was distributed in a bimodal manner (being very high for patients that survived, and negative for those that died), this parameter could not be used in regression analyses that assume a Gaussian distribution of the dependent variable. When attempting to use the net-benefit parameter, this resulted in negative variances in the fitting process. In interventions that use continuous rather than

dichotomous outcomes, the use of net-benefit is likely to be more feasible. An alternative approach in the future would be the use of a multivariate response model where both costs and outcomes are modelled simultaneously and then combined to calculate the ICER (Rice and Jones 1997).

The analysis of cost differences alone implies that this is not a complete economic evaluation, but rather a cost analysis with perhaps lesser decision relevance. An economic evaluation of this trial has been completed (Lubell, Yeung et al. In press, and Annex 3) but no attempt was made at using MLM, and pooled estimates were used for both costs and outcomes as there was no *a priori* plan to stratify the data since no differences in treatment effect were anticipated.

There are a number of other advantages to the use of MLM that were not explored in this study. First is that the presence of missing data does not invalidate the observation (as long as the data are missing randomly). Similarly MLM can be employed even if the data are unbalanced, with differences in sample sizes in the different higher level units (Centre for Multilevel Modelling 2008). Multilevel models have also been shown to be better at predicting outcomes in new sites than both pooled and stratified analyses (Gelman 2005).

7.7 *Chapter conclusion*

The challenges posed in the interpretation of multi-centre trials are considerable. In this chapter the use of MLM for estimating cost differences between two antimalarials in different trial sites was explored, and found to be potentially decisive in the inferences made. Considering the results for each site individually resulted in very high variability between the sites and considerable uncertainty around the estimates. Pooling the data resulted in artesunate being significantly more expensive than quinine. When using MLM to account for the variability between sites, the mean difference was found to be smaller and no longer statistically significant, indicating that artesunate may in fact be less costly.

While this study, with its focus on costs alone, was mostly carried out for illustrative purposes, this information could be of high interest to policy makers reluctant to switch to artesunate for cost considerations, despite the evidence suggesting it to be clinically superior. The cost-effectiveness of artesunate for the treatment of malaria was found to be in the range of $140 per death averted, a highly efficient intervention by any standard. The analysis here confirmed that not only can the use of artesunate be cost-effective, it is in fact not as costly as would appear in standard analyses, with a small chance of it being cost saving, a result that only became apparent when MLM was used instead of standard regression.

8. Markov model for the evaluation of ACTs in the Ugandan HMM programme

As evident in previous chapters and the literature review, decision trees are the most commonly used structure in economic evaluation to represent a health problem and the alternative interventions to address it. For a number of reasons, however, the evaluation of interventions and strategies in the context of malaria requires a more complex decision model. This is particularly true in the case of Home Management of Malaria (HMM) programmes that address recurrent events (febrile episodes) where patients can repeatedly transit between different health states over a long period of time.

Decision trees are most appropriate to represent events that occur in a linear fashion. Constructing the tree to allow for a return to previous states rapidly results in unwieldy structures. Such a structure would become further convoluted where the choice of action in one event (e.g. provision of an antimalarial to a febrile patient) influences the probability and outcome for subsequent ones, for example by providing a prophylactic effect. In addition decision trees do not account for temporality, and so are less suited to evaluating costs and benefits that occur along the progression path with respect to the amount of time spent in each condition prior to transiting to other states.

Markov models offer an alternative structure better suited to address these characteristics. As reviewed in Chapter 4, Markov models' basic building blocks are a predefined number of mutually exclusive health states, through which a patient, or more often a simulated cohort of patients, will transition in a series of cycles (Sonnenberg and Beck 1993). The proportion of the cohort that shifts to each state is determined by transition probabilities which determine how likely a patient is to either remain in the state, or transit to a different one.

Each one of the states is assigned attributes such as costs and quality of life. These values, along with the number of patients in each state, determine the total costs and health

consequences in each cycle. The model can then be run for either a predetermined number of cycles, or until all patients enter an absorbing state (most commonly death). By running this process for the alternative interventions, and summing all costs and consequences during all cycles for each of these, a comparison can be made between their costs and consequences to identify the most efficient option.

In this analysis a Markov model is used to assess the cost-effectiveness of the Ugandan HMM programme, with either an ACT or SP+CQ being distributed, and using a baseline comparator of no HMM programme.

8.1 Background

Following the Abuja commitment to ensure prompt provision of effective antimalarials to 60% of childhood fevers (World Health Organization 2003), eighteen countries in SSA have so far implemented home management of malaria (HMM) programmes (Ajayi, Browne et al. 2008). While the design of these programmes varies, the overall aim is shared, namely bringing effective antimalarials closer to communities for rapid access whenever a child has a fever suggestive of malaria. As the drive continues to introduce ACTs for first line treatment of malaria, in place of failing antimalarials, the use of ACTs is being considered in many HMM programmes as well as in health care facilities (Uganda Ministry of Health 2005; Ajayi, Browne et al. 2008).

Despite the high levels of support for the implementation of HMM in SSA, as evident in the RBM strategies (WHO 2004), so far there is only limited evidence to endorse this strategy (Kidane and Morrow 2000; Sirima, Konate et al. 2003). A recent review of studies on HMM systems in SSA did not find conclusive evidence of their effectiveness, although the drugs, delivery systems and settings in the studies differed considerably, as did methods used to evaluate them (Hopkins, Talisuna et al. 2007). None of the studies identified in the review had included ACTs in the programme.

8.1.1 The Ugandan HMM programme

Since 2002, Uganda has stepped up its efforts to meet the Abuja declaration aim of ensuring rapid access to antimalarials, with the implementation of an HMM programme across the country (Fapohunda, Plowman et al. 2004). The primary distribution mechanism for the antimalarials are community drug distributors (CDDs). Initially the HMM programme in Uganda used a combination of chloroquine and sulphadoxine-pyrimethamine (HOMAPAK®), although artemether-lumefantrine (AL) is now being gradually rolled out in the HMM programme, as well as in health facilities where it is already recommended for first line treatment (Uganda Ministry of Health 2005).

The CDDs are volunteers who are trained to recognize malaria symptoms and danger signs of severe malaria, and either provide carers with effective antimalarials to use at home, or refer the patient to a health care facility. CDDs are not formally paid for their work, although the Uganda guidelines on implementation of HMM do state that reducing the attrition rate can be achieved through provision of financial and other incentives by local governments (Uganda Ministry of Health 2005).

These individuals collect the drugs from central storage sites and store them in their homes. There are approximately 2 CDDs assigned to each village. The CDDs are responsible for replenishing drug supplies in their homes from regional storage sites as these are used up. There is no official reimbursement mechanism in place for this. The CDDs receive 2 weeks of training for their responsibilities as a whole, and a 2 day workshop relating to HMM and ACTs in particular. Their duties as described in the MoH guidelines on training CDDs are listed in Box 8-1.

A number of studies have been carried out in Uganda to evaluate the effectiveness of HMM in increasing access to antimalarials, and their acceptability in the community, showing that the programme has improved access to antimalarials (Fapohunda, Plowman et al. 2004; Ajayi et al. 2008; Nsungwa-Sabiiti, Peterson et al. 2007). However, no studies have been carried out to evaluate the impact of the MoH HMM programme on health outcomes.

- Treating children who have fever/malaria

- Educating mothers on the need for prompt treatment and compliance with the treatment regimen

- Identifying children who need to be referred to the health facilities and advising the caretakers on the need for referral

- Following up treated children to ensure that they comply with the treatment and advice

- Recording treatment given, its outcome and reporting to the nearest health facility

- Working with the community to collect the medicines from the nearest health facility or distribution centre

Box 8-1: Responsibilities of the CDD in Uganda's HMM programme (Taken from the MoH guidelines (Uganda Ministry of Health 2005))

Recently, a trial evaluating the effectiveness of a small scale HMM programme was conducted in Kampala, with the intent of exploring the health benefits of an HMM strategy using an ACT and a highly effective distribution mechanism (Staedke et al. In press). The baseline comparator was standard care, where carers were asked to treat fevers as they normally would in the absence of HMM support. Results for this study were mixed and can be briefly summarised as follows:

1. Access to effective antimalarials for febrile children was superior in the HMM arm

2. There was significantly more over-treatment of non-malarial febrile illness in the HMM arm than in standard care

3. Modest clinical benefit was seen in the HMM arm in the form of lower parasitaemia, and a small and non-significant benefit in reduced hospital admissions

4. The cost of the delivery system for HMM was considerable.

Despite the highly effective drug delivery system used in the trial, the benefits in the HMM arm in terms of health outcome were few. From an economic perspective, when an intervention is found to be more costly and equally or less effective than its comparator, it is considered to be *dominated* with no justification for its adoption. In this instance such a conclusion may not be warranted due to the unique trial circumstances. Firstly, the cost the distribution system was far higher than that of the MoH programme. Secondly, the relatively low prevalence of malaria in the area implied less potential benefit from implementing an HMM programme. Lastly, with good access to health services in their urban environment (including the adjacent Mulago Hospital), children were more likely to receive accurate diagnoses and appropriate treatment for both malaria and non-malarial febrile illnesses, and they were less likely to succumb to severe illness. Nevertheless, these results urge caution before full scale implementation of ACTs in HMM, particularly in urban areas.

8.1.2 Considerations for policy regarding use of ACTs in HMM

Despite the promise of effective treatment for malaria cases, the incorporation of ACTs into HMM programmes raises the same concerns as in the context of their use in health care facilities, primarily regarding their higher costs, risk of adverse effects, and the possibility of parasites developing resistance to the drugs. In the context of HMM, presumptive treatment with ACT can be of even greater concern as a higher degree of malaria over-diagnosis than that already observed in health facilities can be expected (Reyburn, Mbatia et al. 2004; D'Alessandro, Talisuna et al. 2005). Consequently, the provision of antimalarials in all cases of fever may delay care seeking for other non-malarial febrile illness (NMFI), which can be associated with higher mortality as is the case of pneumonia (Kallander, Hildenwall et al. 2008).

The decision therefore that many policy makers in SSA are facing is whether to roll out ACTs in HMM systems, a strategy that Uganda, for instance, is in the process of adopting, despite the lack of evidence to support its effectiveness. More broadly, however, HMM programmes as a whole have never been subjected to an economic evaluation. It is

unlikely however that there will be a single answer to whether HMM is in fact an efficient strategy, as this would most likely be determined by a number of locally specific factors.

The choice of antimalarial is the first and most obvious factor that will determine the cost-effectiveness of HMM programmes, in response to the antimalarial's cost, effectiveness, and possible prophylactic effect. The endemicity of malaria will also be highly influential, in that in areas where this is low, febrile illnesses are most likely to have alternative causes. Presumptive treatment with antimalarials in such settings might not only be ineffective and inefficient, but can also delay seeking alternative treatment for the real cause of illness.

Lastly, the degree to which alternative care is available will also influence the cost-effectiveness of HMM. While HMM is likely to be beneficial where access to health care is poor, where high quality care from health facilities is available, both malaria and non-malarial febrile illness can be effectively diagnosed and treated (Shillcutt, Morel et al. 2008). This can potentially render an HMM programme less advantageous. Similarly, the benefits of HMM will be greater in areas where inpatient care for severe illness is not readily available, as opposed to urban areas where access to higher level health facilities should reduce the probability of severe consequences of malaria and NMFIs.

This chapter uses a Markov model to identify the circumstances under which HMM is appropriate with respect to these factors. Use of the model is demonstrated with data from Uganda, although the model was developed as a DST that can be adapted to other settings.

8.2 Methods

8.2.1 Evaluation framework

The economic framework the model uses is a cost-benefit analysis to calculate the benefit-cost ratio (BCR) for implementation of HMM compared to standard care. Results are presented across a range of malaria incidence levels and degrees of access to care. Health outcomes are measured using the number of years of life lost (YLLs), with monetary values attached in accordance with the WHO benchmark of $150 as applied to DALYs (WHO 2006c). YLLs were calculated for each arm across all cycles. The number of YLLs for each death was derived from the relevant life expectancy table, in this instance that for Uganda in the year 2007 (United Nations Population Division 2005) and discounted at 3% as recommended in the WHO-CHOICE framework (Edejer, Baltussen et al. 2003).

The BCR is specified numerically in instances where HMM is more costly and more effective than standard care. In such instances the BCR can be above 1, indicating that the use of HMM has incremental benefits that outweigh its incremental costs, or below 1 indicating higher incremental costs than benefits. The other possible configurations of differences in costs and benefits were specified with the use of categories rather than specifying a numeric BCR. This was done primarily because a numeric value is of less interest where, for instance, the costs of HMM are higher and the effectiveness is lower. This also avoided ambiguity around the interpretation of the ratio, where a positive value could be a result of higher costs and effectiveness, or lower costs and effectiveness. The possible outcomes are, therefore, as follows:

1. HMM *dominates*, i.e. is more effective and less costly

2. HMM is more effective and more costly, and the incremental benefit is higher than the incremental cost, therefore the BCR is above 1

3. HMM is more effective and more costly, although with a monetary value for benefits that is less than the incremental cost, therefore the BCR is positive but has a value between 0 and 1

4. HMM is less costly but less effective

5. HMM is *dominated* implying that it is both less effective and more costly than standard treatment

From a decision maker's perspective, the first two outcomes would justify the adoption of the intervention unequivocally, subject to budget constraints. Outcome 3 suggests that although HMM is more effective, its incremental cost is higher than the incremental gain, therefore its adoption is not justified. Outcome 4, where both costs and effectiveness are lower for HMM, can be more ambiguous. Although there could be economic gains in implementing such an intervention, since an alternative use of the resources saved might provide greater health returns, there is a strong ethical objection to switching from a more effective strategy to a less effective one. Outcome 5 would suggest that HMM should be rejected.

This classification is taken from the CEA/CUA literature (Drummond, O'Brien et al. 2005), where costs and consequences are treated as separate entities. In the example in this chapter, while the BCR is useful as a summary measure where the benefits of HMM outweigh the costs, the more complex classification provides a clearer indication of where the use of HMM is more or less advantageous on either cost or health outcomes basis alone.

8.2.2 Model structure

The model assumes patients can be in one of 9 mutually exclusive health states, as depicted in Figure 8-1. The probabilities for transition between these states are affected by the prevalence of malaria, the efficacy and prophylactic effect of the antimalarials used, the estimated number of non-malarial febrile illnesses per year, and case-fatality rates for severe illness.
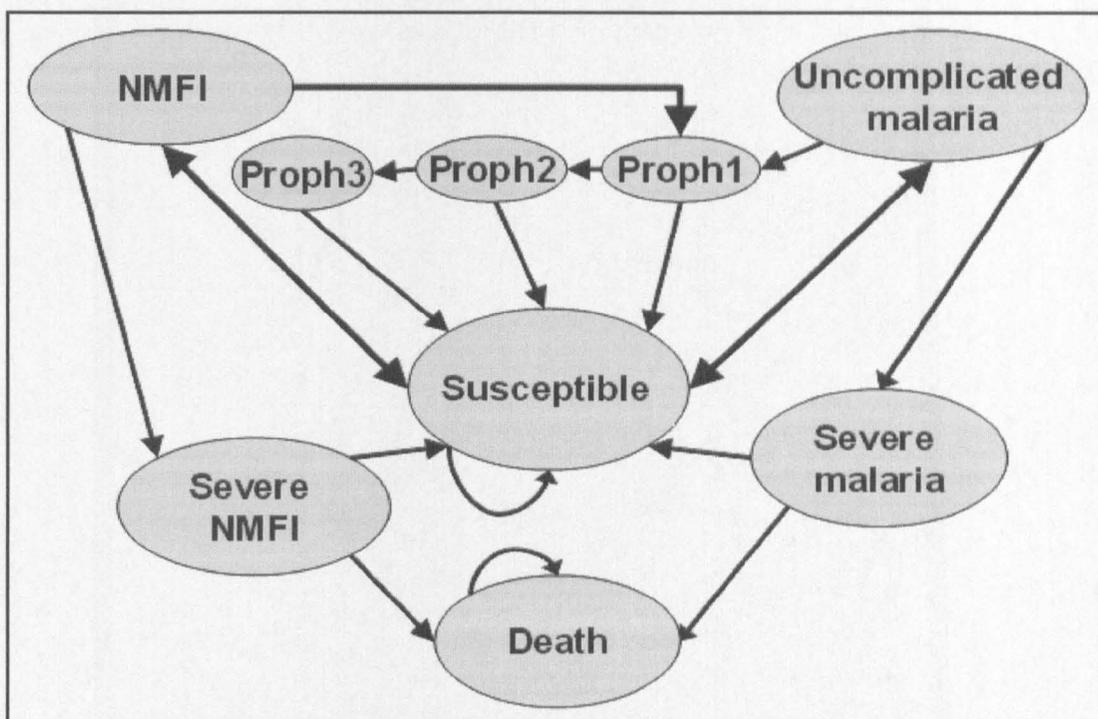
Figure 8-1: Markov health states and possible transitions between them. NMFI - Non malarial febrile illness; Proph – prophylactic effect (followed by number of weeks)

Three of these states are 'tunnel states', where the patient benefits from a prophylactic effect of the antimalarial, however a patient can only remain in this state for one cycle. A child can benefit from a prophylactic effect after receiving an antimalarial regardless of their actual cause of fever, therefore they may transition there from both *uncomplicated malaria* and *NMFI* states. If however a child with an NMFI does not receive presumptive treatment they will return to the *susceptible* state, unless they had progressed to severe illness. The tunnel effect circumvents one of the main limitations of Markov models, that the model has no 'memory' of where an individual patient was in a preceding cycle (Sonnenberg and Beck, 1993). Use of the tunnel states allows for the model to determine how many weeks a patient benefits from the prophylactic effect of the drug, based on the antimalarial being evaluated.

The model is run using weekly cycles, where during each week a cohort of 1000 patients transits through the different health states, with the proportion in each state determined

172

by the transition probabilities. This is repeated to simulate the transitions over 5 years, the number of years for which children are targeted for treatment in the HMM programme.

8.2.3 Input parameters

The following input parameters were used in constructing the model (details on their estimation for the Ugandan HMM programme are provided below):

- Cost of HMM distribution per child per month

- Cost, effectiveness, and prophylactic effect of the drugs used in the HMM arm and those used in health facilities

- Costs of outpatient care for patients with uncomplicated malaria in health facilities

- Cost of inpatient care for patients with severe malaria and severe NMFI

- Patient expenditure for both HMM and standard care

- Transition probabilities between the different Markov health states for patients receiving effective treatment (antimalarials for malaria cases, antibiotic for bacterial NMFIs)

- Transition probabilities for patients with malaria and NMFIs that do not receive appropriate treatment

- Number of NMFIs per year and the proportion of NMFIs that are bacterial

- The value assigned to a year of life lost (or the equivalent loss due to disability).

These inputs are combined with the Markov states to define the transition probabilities between the states and the cost and outcomes associated with them.

8.2.4 Perspective of the analysis

The model is designed to facilitate the comparison of interventions from a number of perspectives. The provider perspective uses the provider costs for HMM distribution, drugs, OPDs and inpatient care, and is combined with the value of YLLs averted through each arm. The patient perspective uses the same measure of effectiveness, but only patient expenditure. The societal perspective combines the costs of patients and providers, and adds a cost in the form of the harm of treatment, as discussed in Chapter 5.

8.2.5 Model output

The comparison between HMM and standard care is made across different levels of malaria incidence and varying degrees of access to high quality health facilities, as shown in Figure 8-2.

Malaria Incidence. Results are presented across a range of incidence levels, reflecting the available estimates of incidence for children under 5 years of age in malaria endemic areas (Roca-Feltrer et al. 2008; Fapohunda et al. 2004). These were reported as being 0.1 to 1.9 per child per annum.

Access to health facilities. This measure is an abstraction and is not based on actual data. The different levels are defined by the proportion of patients who receive high quality care, i.e. correct diagnosis and the recommended first line treatment for malaria cases, or antibiotics for bacterial infection, as opposed to those who do not receive appropriate treatment (i.e. facilities with no diagnostic capacity and the availability of only failing antimalarials). At a level of 20%, for instance, the transition probabilities for 80% of the patients in the model would be based on the estimates of health outcomes for malaria and other illness without appropriate treatment (Table 8-5). The other 20% of patients would have different transition probabilities, based on the estimates of health outcomes for patients receiving correct diagnoses and treatment at health facilities.

The model provides output in the form of a chart indicating where HMM is likely to be cost-effective, in relation to malaria incidence and the availability of high quality health facilities. The chart indicates the results given different configurations of these two factors.

174

	Adjust other model parameters				Perspective:	Provider ▼
					Value YLL:	360 ▼

	Drug	Cost	Effectiveness	Prophylactic (weeks)		
1	Amodiaquine	$0.5	80%	2		
2	SP+CQ	$0.2	75%	1	HMM Drug	CQ ▼
3	AL	$0.9	92%	0.5	Health facility Drug	AL ▼
4	Piperaquine	$0.5	90%	3		
5	CQ	$0.1	60%	10		
6	Custom 1	$4.0	60%	0.1		
7	Custom 2	$0.1	100%	6		

Malaria incidence per child per year

Access to health facilities	0.1	0.5	1.0	1.5	2.0
100%	LCLF	LCLF	LCLF	LCLF	LCLF
80%	LCLF	LCLF	LCLF	LCLF	DTS
60%	LCLF	LCLF	DTS	DTS	DTS
40%	LCLF	DTS	DTS	DTS	DTS
20%	3.4	4568.9	DTS	DTS	DTS
0%	78.8	307.5	471.5	572.7	651.9

Key:
DTS	HMM dominates
25.0	Benefit cost ratio above 1
0.5	Benefit cost ratio below 1
LCLF	HMM less costly but less effective
	HMM dominated

Figure 8-2: Model interface. The lower panel is the model output indicating the circumstances in which HMM is likely to be appropriate. Above this are the controls where costs and transition probabilities can be adjusted and drugs can be chosen for both HMM and health facilities.

The model was designed as a decision support tool and allowed for the incorporation of data from different settings, choice of antimalarial to be used in HMM and in health facilities, and for different estimates of the transition probabilities. In this analysis the data used related to the Uganda HMM programme, evaluating HMM using either HOMAPAK® or AL, as compared to standard care with the use of AL in health facilities.

8.2.6 Data used in the analysis of the Uganda HMM programme

Costs of the Uganda HMM programme. An overview of the HMM programme and its costs were obtained in an interview with Dr Fred Kato of the Malaria Control Programme, who also provided a number of MoH documents detailing the expenditure on the programme. These relate to the training of CDDs, their monitoring, and the cost of antimalarials supplied. The distribution system is based primarily on the work of the CDDs and although they are volunteers, there is an opportunity cost to their time spent on this intervention. The MoH data were supplemented therefore by interviews with three CDDs to obtain an

indication of their own time and expenditure on HMM activities. CDD time was then assigned a monetary value equivalent to the Ugandan average wage (Drummond et al. 2005; ILO 2008).

Costs of high quality outpatient care: A comprehensive estimate for the cost of high quality outpatient care was obtained by identifying a recently constructed clinic in the vicinity of Jinja, western Uganda, that maintained high quality diagnostic and treatment capacities. The clinic was purposely chosen as likely to be at the highest end of what can be reasonably expected in the context of outpatient care in SSA. Costs were obtained for construction, overheads and variable inputs. Capital costs were annuitized over 5 years to obtain a monthly equivalent cost. The average costs per OP visit are dependent on the volume of patients seen, so records of patient visits over the previous 3 months were sought to assess the volume. The clinic was reported to be running far below capacity so the cost per patient was modified to an 80% level of utilization, as per WHO recommendations (Edejer et al. 2003).

The cost of the antimalarials provided in health facilities was taken to be that of AL, the current first line treatment for malaria in Uganda.

Cost of inpatient care for severe malaria and non-malarial febrile illness: A mid-sized hospital in Kisiizi, Southwest Uganda, was chosen to calculate the costs for treating children with severe illness. The costing process combined micro costing, using patient records for treatments received and labour costs for clinicians and nurses on the ward, with step-down hospital costing to assess the expenditure of service departments and the proportion of these dedicated to the paediatric ward. Hospital capital costs were obtained from accounting records showing their estimated present value. These were annuitized to obtain a monthly equivalent. Average length of stay was computed from patient records for both malaria and NMFIs. The hospital was functioning at full capacity so the overall costs were assigned to the number of patient days per month without adjustments for under-utilization.

Patient costs. Patient costs were obtained from the trial carried out in Kampala for both HMM and standard care (Staedke et al., in press). In the trial, patients were asked to record any expenses related to the management of febrile illness. These included user fees, travel expenses, drugs purchased and any other illness related expenditure. Productivity losses in the form of time spent caring for children with febrile episodes were added based on the time spent caring for children multiplied by the average wage for unskilled workers (Drummond et al. 2005; ILO 2008).

Transition probabilities. The probability of contracting malaria is determined by the incidence, estimates for which were obtained as described above. For malaria patients who are correctly diagnosed and treated the transition probabilities were taken from the literature estimating the effectiveness of AL and HOMAPAK® in Uganda (Staedke, Kamya et al. 2001; Obua, Gustafsson et al. 2006; Yeka, Dorsey et al. 2008).

The probability of developing an NMFI was based on estimates for the incidence of febrile episodes per year (Breman 2001), and subtracting the incidence of malarial episodes.

Transition probabilities for patients with no access to care or incorrectly treated were obtained using expert opinion on expected health outcomes for patients with untreated malarial and non-malarial febrile illness. For non-malarial febrile illness, expert opinion was sought on the proportional breakdown of these into viral and bacterial illness, along with the transition probabilities for these to severe illness and death. These estimates are provisional as they are taken from the first round of a Delphi survey that aims to collect a wider range of expert opinion in a systematic manner. The protocol and questionnaire for this are presented in Annex 2.

Prophylactic effects. Estimates for the prophylactic effect of the antimalarials were taken from the literature (White 2005). These were used to determine the number of weekly cycles, following receipt of an antimalarial, during which the child is protected from re-infection.

The probability of remaining in a Markov state or returning to the *susceptible* state, where this is possible, is determined by subtracting the probabilities that the patient leaves the

state through other possible paths, such as developing severe illness, from 100%. The parameter estimates and sources used are summarised in Table 8-1.

Costs	Estimate	Source	Notes
HMM distribution	$0.2 per child/month	Primary data, Uganda MoH documents	
OPD care (excluding drugs)	$4.5	Primary data – Jinja clinic	
Inpatient care for severe malaria	$20	Primary data- Kisiizi Hospital	Based on average length of stay
Inpatient care for non-malaria severe illness	$12	Primary data- Kisiizi Hospital	Based on average length of stay
Antimalarial costs	HOMAPAK - $0.15 per child/month AL - $0.65 per child/month	Uganda MoH	
Antibiotic costs	$0.3	Lubell et al. (2008a)	
Transition probabilities			
HOMAPAK 28 failure rate	35%	Staedke et al. (2001)	Drug effectiveness measured by 28 day outcome without genotyping
AL 28 day failure rate	17.3%	Yeka et. al (2008)	
Untreated malaria becoming severe	30%	Preliminary Delphi survey results	These results are from the first round of the survey and are likely to change in the following round as the questionnaire is modified.
CFR untreated severe malaria	62%		
Proportion of NMFI that require antibiotics	30%		
Untreated bacterial NMFI becomes severe	40%		
CFR untreated severe NMFI	30%		
Prophylactic effects AL	3 days	White (2005); Expert opinion (Chris Whitty, Sarah Staedke)	
Prophylactic effect HOMAPAK	1 week		
Life expectancy	52 years	UNPD (2005)	Averaged for age 1 to 5
Discount rate	3%	Gold et al. (1996)	

Table 8-1: Parameter estimates used in the model

8.2.7 Sensitivity analysis

A number of sensitivity analyses were carried out to assess the impact of using different antimalarials in both the HMM programme and in health facilities. The drugs were the most common antimalarials available and estimates were used for their costs, efficacies and prophylactic effects. An analysis was also run using a stylized drug representing an ideal antimalarial, with a cost equivalent to that of CQ, effectiveness of 100%, and a prophylactic effect of 3 weeks.

Results were initially obtained using the provider's perspective. In the sensitivity analysis, first the provider's perspective was broadened to include the harm of treatment factor as discussed in Chapter 5. The model was then run using the societal perspective that combined costs for patients, providers, and the harm of treatment factor.

A sensitivity analysis was also carried out to guage the impact of using a 6% discount rate, as recommended by the WHO-CHOICE (Edejer, Baltussen et al. 2003).

The model was designed using Microso Excel® 2002 and macros were wri en with Microso Visual Basic® 6.3.

8.3 Results

8.3.1 Costs

The costs for the HMM programme as currently run by the MoH and with the use of AL are summarised in Table 8-2, showing that the largest expense is that for the drugs themselves, even with the current use of HOMAPAK®. Other expenses for the MoH are less than $0.1 per child per month. The CDDs, who receive no formal payment, shoulder a quarter of the total estimated economic costs for the programme when HOMAPAK® is the treatment. Substituting this with AL raises the total cost per child considerably, with the cost of the drugs themselves constituting 92% of the programme cost.

	HOMAPAK®			AL		
	Monthly	Annual	% of total	Monthly	Annual	% of total
CDDs	$0.05	$0.60	24%	$0.05	$0.60	7%
MoH (excluding drugs)	$0.01	$0.12	5%	$0.01	$0.12	1%
Drugs	$0.15	$1.80	71%	$0.65	$7.80	92%
Total cost/child	$0.21	$2.52	100%	$0.71	$8.52	100%

Table 8-2: Cost per child of the Uganda HMM programme using HOMAPAK® or AL. CDD - Community drug distributor; MoH - Ministry of Health

The operational costs for a high quality outpatient clinic are summarised in Table 8-3. The cost per patient was estimated by averaging the volume per month over the preceding 3 months, equalling 332 patients, providing a cost per patient of $7.2. This was adjusted by increasing the volume by 30% to reach the 80% recommended utilization rate (based on the opinion of the clinic manager regarding the clinic's full capacity), giving an average cost of $4.5 per patient.

	Monthly	% of total
Salaries	$1,466	61%
Perishables	$460	19%
Utilities, travel, maintenance	$306	13%
Total recurrent	$2,232	93%
Capital	$174	7%
Total	$2,406	100%

Table 8-3: Total costs of running a high quality outpatient clinic in Jinja, Uganda

The inpatient care costs of treating malaria and NMFIs excluding those for drugs are summarised in Table 8-4. When accounting for drugs and the different average length of stay for malaria and NMFI patients, the average cost per case of malaria was estimated to be $11.2, while the cost for NMFI cases was $20.4. The difference is explained by both the slightly lower length of stay for malaria patients, and the lower drug costs of quinine as opposed to treatments for severe NMFIs.

	Monthly	% of total
Salaries	$3,465	60%
Perishables	$750	13%
Utilities, travel, maintenance	$1,211	21%
Total recurrent	**$5,427**	**94%**
Capital	**$358**	**6%**
Total	**$5,785**	**100%**

Table 8-4: Monthly cost for running the inpatient care paediatric ward in southeast Uganda

Total patient expenditure is comprised of the direct costs associated with treating febrile episodes, and with opportunity costs in the form of time taken off work by carers. Cost for treating febrile episodes was lower in the HMM arm, averaging $0.9, as opposed to $1.5 in the standard care arm. The difference in time spent per febrile episode was minor, being 0.85 days for the HMM arm and 0.96 for the standard care arm. Total cost per febrile episodes were, therefore, $1.9 for patients in the HMM arm and $2.8 for those in standard care.

8.3.2 Model output for the use of HOMAPAK® in HMM

With the use of the above costs and the baseline transition probabilities, the output chart indicates a general diagonal division between the top left area of the graph, areas of high access/low incidence, that favour standard care, to bottom right areas of low access and high malaria incidence that favour HMM.
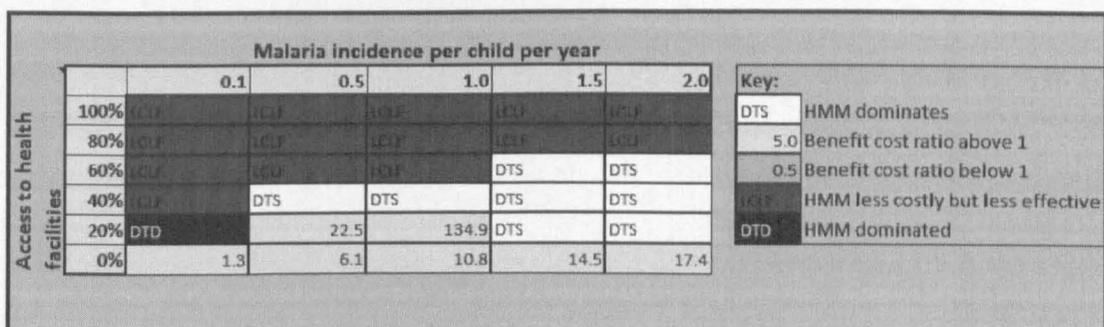
Figure 8-3: Model output for the Uganda HMM programme with the use of HOMAPAK®

Use of HOMAPAK® in the HMM programme appears to be efficient across all incidence levels where there is no access to alternative health care facilities, with BCRs ranging from

1.3 in areas with low incidence and zero access to care, up to 17.4 at a high incidence rate of 2 malaria attacks per child per year (Figure 8-3). In all areas of high access to health facilities, and particularly where transmission is low, HMM is less costly but less effective than standard care. In areas of low incidence HMM is dominated (i.e. is more costly and less effective), where access to care is in the range of 20-40%. Conversely, in these levels of access to care but where malaria incidence is high, HMM dominates standard care.

8.3.3 Model output for the use of AL in HMM

With the use of AL in place of HOMAPAK®, results overall are similar, although HMM is slightly less advantageous particularly in areas of low incidence (Fig 8-4). This is evident in the lower BCRs, which drops below 1 in areas of low incidence and with no access to care.

		Malaria incidence per child per year					Key:	
		0.1	0.5	1.0	1.5	2.0	DTS	HMM dominates
Access to health facilities	100%	LCLF	LCLF	LCLF	LCLF	LCLF	5.0	Benefit cost ratio above 1
	80%	LCLF	LCLF	LCLF	LCLF	LCLF	0.5	Benefit cost ratio below 1
	60%	LCLF	LCLF	DTS	DTS	DTS	LCLF	HMM less costly but less effective
	40%	DTD		3.9	29.0	144.2	DTS	DTD	HMM dominated
	20%	DTD		3.4	7.1	9.9	12.0		
	0%		0.8	3.5	5.9	7.6	8.8		

Figure 8-4: Model output for the Uganda HMM programme with the use of AL

The advantage of HOMAPAK® over AL can be explained as a consequence of both the lower cost of HOMAPAK® and its prolonged prophylactic effect preventing the contraction of further malaria attacks.

8.3.4 Sensitivity analysis

The model was run using a range of estimates for antimalarial costs, efficacies and prophylactic effects, reflecting those of some of the most readily available antimalarials. While these showed some variation in the precise BCRs, there was no substantial impact on results. Even the use of a hypothetical, highly effective and cheap antimalarial in the HMM arm only marginally improved the cost-effectiveness of HMM so that in areas where it already had a positive BCR, this became even higher or dominated standard care (Figure

8-5); however, almost all areas where HMM was not a cost-effective option with the use of AL or HOMAPAK® remained this way also with the use of an ideal antimalarial.

		Malaria incidence per child per year					Key:	
		0.1	0.5	1.0	1.5	2.0	DTS	HMM dominates
Access to health facilities	100%	LCLF	LCLF	LCLF	LCLF	LCLF	5.0	Benefit cost ratio above 1
	80%	LCLF	LCLF	LCLF	LCLF	DTS	0.5	Benefit cost ratio below 1
	60%	LCLF	LCLF	DTS	DTS	DTS	LCLF	HMM less costly but less effective
	40%	LCLF	DTS	DTS	DTS	DTS		
	20%	DTD	DTS	DTS	DTS	DTS	DTD	HMM dominated
	0%	2.1	9.4	16.8	22.6	27.3		

Figure 8-5: Model output for an ideal antimalarial, compared with AL as first line treatment in health facilities

These results, however, did not include the harm of treatment parameter. When introducing this into the analysis, the areas where HMM appeared more beneficial were vastly reduced, particularly in low transmission areas (Figure 8-6). The loss of future lives associated with the extended use of antimalarials under the HMM strategy implies that it is less effective than standard care in all but high incidence areas and where access to health care is very limited.

		Malaria incidence per child per year					Key:	
		0.1	0.5	1.0	1.5	2.0	DTS	HMM dominates
Access to health facilities	100%	LCLF	LCLF	LCLF	LCLF	LCLF	5.0	Benefit cost ratio above 1
	80%	LCLF	LCLF	LCLF	LCLF	LCLF	0.5	Benefit cost ratio below 1
	60%	LCLF	LCLF	LCLF	LCLF	LCLF	LCLF	HMM less costly but less effective
	40%	DTD	DTD	DTD	DTD	LCLF		
	20%	DTD	DTD	DTD		1.5	3.4	DTD	HMM dominated
	0%	DTD	DTD		1.3	2.9	3.9		

Figure 8-6: Model output with the inclusion of the harm of treatment factor, with AL being used in both HMM and health facilities

Using a higher discount rate of 6% had a marginal effect on the outcome, slightly moderating the BCRs where HMM was more costly and more effective than standard treatment. This is due to the costs of treatment being restricted to the 5 years during which the programme operates, while the number of years of life lost when a patient dies were discounted to the full length of the life expectancy of the child, therefore the reduction was proportionally greater.

8.4 Discussion

8.4.1 Assumptions and limitations

A number of simplifying assumptions were made; these are believed to have little or no effect on the difference between the alternative strategies. Firstly the occurrence of neurological sequelae was not accounted for. These were included in the initial iterations but were found to have only negligible impact, while considerably complicating the model given the need for an additional health state to capture them. Secondly, deaths from unrelated causes were ignored in both arms, as the population is of a younger age than is found in most of the literature using Markov models, which relates mostly to older populations with chronic illnesses. Furthermore, the effect of unrelated deaths was assumed to be roughly equal in both arms.

The comparator of standard care was an abstraction comprised of patients who either had access to high quality health facility care, or those that had no access. This is a gross simplification of vast diversity in degrees of access to health services and quality of care. As the aim was to provide an illustration of the variability in HMM cost-effectiveness rather than a precise value, this was assumed to be an adequate representation for these purposes. A more precise assessment of the cost-effectiveness of HMM in particular settings would require further details of the availability of health services, which could then be incorporated in the model.

The model did not account for the development of immunity amongst children less than 5 years of age. While it has been shown that in high transmission areas children might develop partial immunity to severe illness from a very early age (Gupta, Snow et al. 1999),

184

this dynamic is not fully understood and to avoid over complexity in the model this was omitted. The likely effect of including this factor would be to further reduce the benefit of HMM. Inclusion of this factor would also have allowed for an incremental analysis to explore the age groups in which HMM is most cost-effective.

The model excluded the possibility of a child having both an NMFI and malaria simultaneously. With a maximum malaria incidence rate of 2 per year, however, this is not likely to be a very common occurrence, therefore for the sake of simplicity this was not included in the model.

Some of the movements between health states in the model were restricted to avoid the model becoming overly complex. Patients who were in a malaria state, for instance, could not transit straight to an NMFI state, but rather moved back to asymptomatic status, or on to severe malaria. Patients were also assumed to stay in malaria and NMFI states for one cycle, i.e. the length of all febrile episodes was assumed to be one week.

Many of the transition probabilities were derived from expert opinion rather than primary data. As noted, a Delphi survey is being conducted to obtain a broader range of expert opinion; the estimates obtained in this survey will replace the current ones prior to dissemination of the analysis.

The estimates for costs of outpatient and inpatient care were each obtained in a single location. Ideally a greater number of facilities would have been surveyed; however time and budget constraints limited the ability to do this. The estimates do, however, approximate those used in previous analyses in the thesis obtained in Tanzania. Patient costs can differ considerably in different settings. The costs used in this analysis were from an urban setting where transport costs are likely to be lower, as well as the time spent travelling to clinics.

Perhaps the greatest apparent limitation of this model and others of its kind is the multitude of results it provides for the range of configurations available for the variable input factors. This however, should be recognized as the reality of decision making rather than a fault of the methodology.

8.4.2 Use of a Markov model for the evaluation of malaria interventions

Use of the Markov model allowed for more flexible modelling of febrile illness and its management than that available with decision trees. The cyclical nature of febrile illness was captured in the transition between the Markov states. This also allowed for the incorporation of the antimalarials' prophylactic effect in reducing the probability of subsequent malaria episodes and capturing the ongoing costs and effectiveness of the alternative strategies rather than those for a single febrile episode as is possible with simple decision trees.

This structure is critical for the evaluation of HMM programmes, as these address recurrent febrile illnesses in the home. The use of Markov models, however, may also be beneficial in the assessment of treatment in health care facilities where patients will often re-attend following treatment failure, and in other malaria related interventions due to the recurrent nature of the disease and its various chronic manifestations.

The model output provides greater detail than commonly found in economic evaluations where dichotomous outcomes are the norm, for example being above or below a decision threshold. There are, however, a greater number of distinctions that can be useful to policy makers, particularly when results are demonstrated across a range of settings simultaneously. It can be useful to distinguish between situations where, for instance, HMM is more effective and more costly, yet the benefits do not outweigh the costs, and those where HMM is less effective but also less costly to the extent that standard treatment is not cost-effective.

Results were presented mostly from the provider's perspective, as this was assumed to be the initial concern for decision makers. The analysis, however, also examined results from the societal perspective. The inclusion of the potential harm of treatment associated with provision of antimalarials is of particular significance in the context of HMM, given the expected increase in use of antimalarials, which is itself the aim of the intervention. It is imperative, therefore, that due consideration is given to the more comprehensive costs and benefits associated with extensive use of the treatments. In practice, the choice of perspective did not have a significant impact on results.

8.4.3 Policy implications

HMM programmes are being implemented in most malaria endemic countries across SSA in an attempt to ensure that febrile children have rapid access to antimalarials, preventing their possible deterioration to severe illness and death. This analysis has explored the variation in the cost-effectiveness of the strategy as expressed by the BCRs, dependent on transmission intensity and access to health care. The results indicate considerable variation in response to these factors, suggesting that widespread roll out of HMM programmes might lead to misuse of resources, and in some instances worse health outcomes than would otherwise occur. As urbanization brings improved access to health care, and with growing evidence of reduction in incidence of malaria (Gosling et al. 2008; Greenwood et al. 2008), these results are of potentially high significance.

The modelling suggested that in many settings HMM was likely to be less effective, but also less costly, than providing high quality facility based care. This would suggest that in areas where neither services are widely available, HMM might indeed be a good stop-gap. In the long run, however, investment in facility based care is likely to provide better health returns than continual reliance on HMM for the management of childhood fevers.

The model also indicated that the advantages of using AL instead of HOMAPAK® in the HMM programme were modest. This can be explained by the extended prophylactic effect of HOMAPAK®, modifying the advantage of the higher efficacy of AL, and by its lower cost.

While this analysis suggests that the application of HMM programmes in many settings may not be economically justified, much of the advantage of health care facilities over HMM emanates from the greater diagnostic capacity, allowing clinicians to distinguish between non-malarial and malarial illness and treat patients accordingly (Shillcutt, Morel et al. 2008). This capacity could be introduced into HMM programmes with the use of RDTs administered by, for instance, CDDs. This appears to be the logical next step for HMM programmes, with a focus on treating fevers as a whole and not solely the provision of antimalarials (Yeung 2008; Bell and Perkins 2008). However, data are needed to support such strategies.

8.5 Chapter conclusion

To ensure that evaluations of interventions such as HMM programmes provide appropriate decision recommendations, they should adopt model structures that best represent the disease they address. The advantage of using a Markov model is that patients can transit between the Markov states with the relevant costs and outcomes over a long period of time, during which children experienced multiple febrile episodes, both malaria and NMFIs.

HMM programmes are being heavily promoted across SSA to improve access to effective antimalarials. There are likely to be benefits to this, particularly in areas where malaria is rife and where health facilities are not readily accessible. At the same time, as this analysis has indicated, there are many instances where the use of HMM programmes may not effective or efficient. Promoting the strategy where malaria incidence is low, and where alternative care with better diagnostic capacity is available may lead to extended overdiagnosis of malaria, missing other causes of illness, and further encouraging the development of resistance to antimalarials.

9. Discussion

In the opening chapter this thesis presented different approaches to decision making in health care in general, and malaria in particular, arguing for the need for analytic and explicit methods to weigh up the multitude of costs and benefits associated with malaria treatment and diagnosis. Two basic themes ran through subsequent analyses in the thesis: (1) the need to capture the broader and long term impacts of malaria and the interventions used to control it; (2) the need to ensure that decision recommendations are relevant to the time and place in which they are being used. To realize these needs, methods and models new to the context of the evaluation of malaria control were employed. Figure 10.1 shows how the themes and analyses tie together in the thesis as a whole.
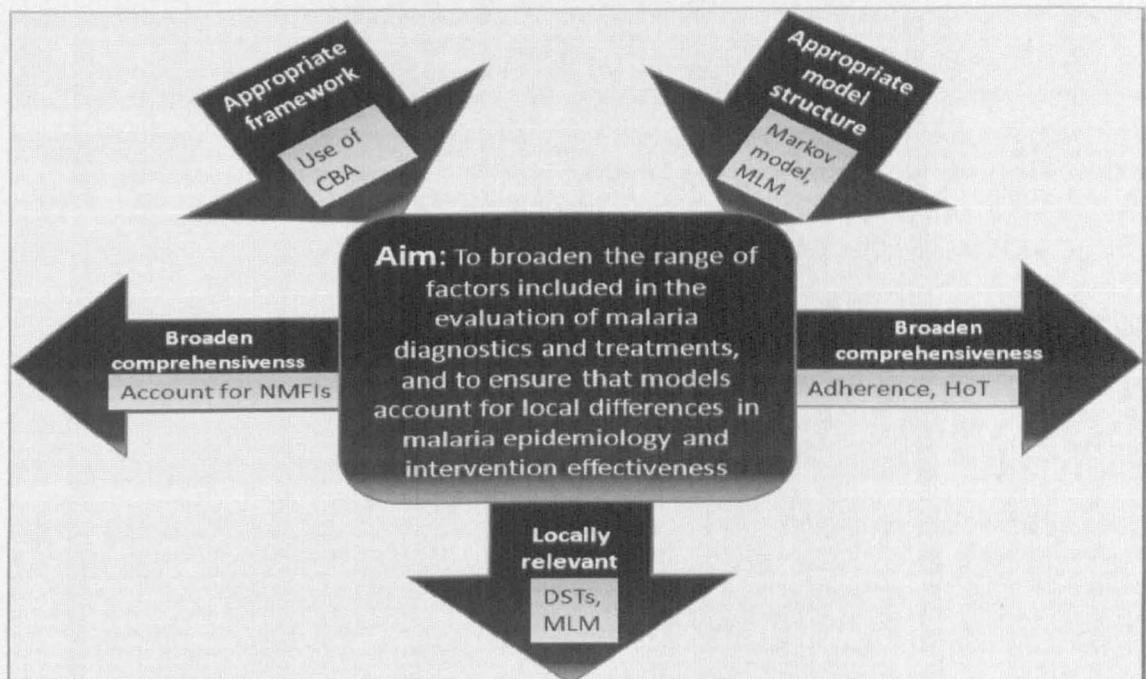
Figure 9-1: An illustration of the thesis aim, themes and how these were expressed in the analyses. CBA – Cost-benefit Analysis; MLM – Multilevel Model; NMFI – Non-Malarial Febrile Illness; HoT – Harm of Treatment; DST – Decision Support Tool

The analyses in the thesis have explored these themes, focusing on different areas as required by the policy question being raised.

While each of the analyses has been discussed individually throughout, there are a number of general points that relate to the gaps identified in the literature, and the means by which this thesis has attempted to address them. The discussion of these issues is followed by a brief summary of the main findings and policy implications. Lastly some of the areas identified for further research are presented.

9.1 Gaps in the literature and strengths and weaknesses of the methods used to address these

The literature review in Chapter 2 examined the methods used in previous economic evaluations of malaria diagnostics and treatments. Although a small number of these evaluations did employ appropriate frameworks and advanced modelling techniques, on the whole, there were multiple shared weaknesses that reduced their relevance for decision making purposes.

9.1.1 Choice of economic evaluation framework

Almost all previous evaluations of malaria control interventions employed either cost-effectiveness or cost-utility analyses. The limitation of cost-effectiveness analyses using disease-specific measures of outcome was one of the shortcomings of many previous evaluations. For instance, evaluations of malaria diagnostic tests have used the cost per case detected as a measure of outcome (Bualombai, Prajakwon et al. 2003; Rolland, Checci et al. 2006; Fernando, Karunawee et al. 2004). Similarly, evaluations of antimalarials have looked at cost per patient treated (Sudre, Breman et al 1992; Gogtay, Kadam et al. 2003; Chanda, Masiye et al. 2007). Other examples of such studies are those for bed-nets and internal residual spraying, using measures of outcome such as cost per child protected (Guyatt, Kinnear et al. 2002), or the cost per bed-net distributed (Stevens, Wiseman et al. 2005). The limitation of this framework is that decision makers rarely have any frame of

reference when considering these results, as they are not comparable to each other or to other interventions.

Cost-utility analyses partly resolve this by offering generic measures of outcome, allowing for comparison of interventions with different aims. However, they still require a decision threshold to indicate whether the intervention is cost-effective. The most comprehensive studies are those by Goodman et al (2000) and Morel et al. (2005), which evaluated a wide range of malaria control measures, with respect to the cost per DALY averted and the WHO threshold of $25 and $150 per DALY averted. Shillcutt et al. (2008) used a probabilistic sensitivity analysis to juxtapose a range of costs per DALY averted against different decision thresholds to indicate where the use of RDTs is likely to be cost-effective.

As both DALYs and QALYs are continuously being refined, so are the methods for assessing the monetary value of averting the loss of a DALY or gaining a QALY (Sachs 2002; Shillcutt, Walker et al. Unpublished). However, the incorporation of such measures into the evaluations has led these back in to the realm of cost-benefit analysis, even though many analysts continue to skirt around the explicit valuation of health outcome in monetary terms (Gafni and Birch 2006).

The use of a cost-benefit framework facilitates more coherent analyses and outcomes, as evident in the studies in this thesis. One key advantage is the ability to incorporate a greater range of factors beyond immediate costs and benefits into the analysis in a more coherent fashion. This was demonstrated with the incorporation of the harm of treatment factor, where a monetary cost was assigned to the potential adverse outcomes associated with the use of antimalarials, which can be of high significance under strategies of presumptive treatment and HMM. The presumptive treatment of children in many low and medium transmission areas, for instance, no longer appeared efficient once the harm of treatment costs were included in the analysis, as shown in Chapter 5.

Similarly, as was discussed in Chapter 7, the use of net-benefit as a measure of outcome (approximating a CBA) in regression analyses is a particularly useful method to explore the

factors that determine an intervention's efficiency (Hoch, Briggs et al 2002). The use of CBAs combined with regression analysis represents a potentially superior framework for trial based evaluations in general, and for multi-centre trials in particular, with the use of multilevel modelling (Manca, Rice et al. 2005). This does, however, require that health outcomes are valued in monetary terms, so that the dependent variable can capture both costs and outcomes simultaneously.

9.1.2 Comprehensiveness of the analysis

Many previous evaluations have chosen to focus on a limited number of parameters, notably those which are more readily available for collection in clinical trials, such as immediate outcomes measured in natural units and provider costs (Bualombai, Prajakwon et al. 2003; Rolland, Checci et al. 2006; Fernando, Karunawee et al. 2004). Other factors, however, have been shown to have a critical impact on the efficiency of interventions, and their inclusion in the analysis is therefore imperative, rather than being left as afterthoughts. This was demonstrated in previous evaluations by Yeung (2006) and Coleman et al. (2004), when looking at the impact of parasite resistance to antimalarials on the cost-effectiveness of the treatments.

One factor introduced systematically into the analyses in the thesis is the health outcome for patients with non-malarial febrile illness. This was identified as a key issue in the evaluation of RDTs by Shillcutt et al. (2008). The confounding of malaria and febrile illness has resulted in the disregarding of outcomes for patients with other febrile illnesses who are treated with antimalarials under strategies of presumptive treatment. Analyses that ignore this will overstate the benefits of antimalarials and ignore the reality that a considerable proportion of patients do not benefit from the treatment. Moreover, without appropriate treatment these patients are at higher risk for severe outcome from their true cause of illness (Amexo, Tolhurst et al. 2004). Throughout the studies in this thesis, estimates have been made for the proportion of patients with NMFIs and their health outcomes as well as for patients with malaria.

By expressing the cost-effectiveness of interventions across a range of prevalences of malaria amongst febrile patients, the impact of the proportion of NMFIs on the efficiency of interventions is observable, as shown in Chapters 5 and 6. The main difficulty, however, is in estimating the breakdown of NMFIs into those illnesses that require antibiotic or other treatment, as opposed to those that are likely to be viral and mostly self limiting. Localising these assessments in decision models is likely to provide more adequate recommendations (English and Scott 2008). Lack of data meant that in the analyses here, the estimates for the proportion of NMFIs that are bacterial had to be based on expert opinion. The Delphi survey underway seeks a broader range of expert opinion on this for different age groups and malaria transmission intensities.

Chapter 6 explored how the apparent efficiency of RDTs and ACTs is challenged when analyses are extended beyond a focus on immediately observable costs and consequences, and incorporate non-adherence to test results and the harm associated with provision of antimalarials. No previous studies have accounted for these factors when evaluating diagnostic tests for malaria (or indeed for diagnostic tests for other diseases as far as could be found in a brief survey of the literature), despite the evidence of compromised adherence to RDTs and microscopy and high rates of over-prescription of antimalarials.

Including the adherence factor in the analysis of RDTs indicates firstly the reduced efficiency of rolling out the tests given current prevailing adherence levels, and the potential gains of implementing training programmes to increase adherence. The inclusion of the harm of treatment factor allows policy makers to observe the full benefits of using RDTs in diminishing the unnecessary use of antimalarials.

The main weakness in this greater comprehensiveness is that such factors are difficult to estimate, introducing further uncertainty into the analysis, and in some instances appearing 'unscientific', thus compromising credibility. Furthermore, there will inevitably be a wide range of additional costs and consequences that could be introduced, and to a certain extent the choice of which to include will be subjective or even arbitrary. The adherence factor, for instance, might appear more reliable in its estimates, being directly

measurable, and of immediate relevance in that policy makers can act upon it. The harm of treatment factor, on the other hand, is both more difficult to assess, and its policy implications are less tangible.

The apparent arbitrariness in choice of parameters and the high uncertainty surrounding some of the estimates can indeed raise doubts around the accuracy of the analysis, particularly if the decision on parameters and estimates are not explicitly specified for stakeholders to consider their relevance. This can be partly addressed by presenting a range of results across a plausible range of estimates, as was done in the present analyses. A further advantage of two of the models in this thesis is their user-friendly design that facilitates both the inclusion/exclusion of parameters and their adaptation according to the users' own beliefs and opinions. It is possible that this might increase the confidence stakeholders have in the validity of results, although this was not tested in the thesis.

One limitation with respect to the comprehensiveness of the analyses is that the models did not respond to changes in transmission intensity that might occur following the adoption of the proposed interventions. The main reason this was excluded is that in high transmission areas which comprise much of SSA, no individual intervention is believed to have the ability to significantly impact transmission (Molineaux and Gramiccia, 1980; Lines et al. 2008). Nevertheless, dynamic models that account for transmission changes do indeed have much to offer to the evaluation of malaria diagnostics and treatments in areas of lower endemicity (Yeung 2006). This is an area however that requires considerable research in its own right prior to inclusion in decision models. Two recent doctoral theses at LSHTM have focused on these issues and it is hoped that future work can incorporate these dynamics in the models (Yeung 2006, Okell et al. 2008).

9.1.3 Local relevance of results

Perhaps the greatest potential weakness in the entire approach to economic evaluation is the over-generalisation of results. The incentive to do so exists not only for the analyst, keen to ensure that their work has significance beyond a particular setting, but also for the reader and decision maker, seeking evidence to support new policies or the adoption of

194

new interventions in their own setting. The reality of malaria epidemiology and control, however, is a highly dynamic one with wide regional diversity (Ye, Kyobutungi et al. 2007). Malaria incidence is changing continuously, with growing evidence emerging in the last few years of its overall reduction across SSA (Gosling et al. 2008; Ceesay, Casals-Pascual, et al. 2008). This will have implications not only for numbers of cases, but also for individual immunity to severe illness and death once infected. Furthermore, the number of drug combinations in the ACT class and RDT brands is continuously growing, rendering questions such as 'are RDTs/ACTs cost-effective?' meaningless, without further specification of the specific drug or test in question.

Publications based on data collected at a specific point and time might be uninformative, or indeed misleading when applied elsewhere several years later. In this thesis two analyses focused on methods to ensure that evaluation results are applicable to different settings, the first of which related to modelling based analyses, the second to multi-centre trial based evaluations.

In Chapter 6, decision support tools were constructed for easy adaptation of economic evaluation models by decision makers to their own circumstances. Two alternative frameworks were described for such DSTs, with a trade-off in complexity and methodological validity between the two. In an age of advanced IT capacities and with a growing number of individuals trained in at least a basic understanding of health economics, the use of DSTs in lieu of static evaluations seems like a natural progression for economic evaluation (van Gool, Gallego et al. 2007, English and Scott 2008). The diffusion of decision making powers that characterises much of health care reform could facilitate this further. While there will always be a need for rigorous peer review processes to assess the validity of evaluations, these could be directed at ensuring the quality of the model structure, and allowing local stakeholders to populate the models with local data and their own assumptions and preferences.

One limitation to the possible dissemination of DSTs is the degree to which policy makers are adept in the use of concepts such as probabilities and measures of outcome such as DALYs (Teerawattananon 2007). Similarly, it may be the case that local stakeholders do not

have access to higher quality data than the analyst. For this reason DSTs could initially be populated with best available data, as was done in the models here, which can then be modified as appropriate by local stakeholders and policy makers.

The main limitation may not be the technical skills required to manage the DSTs, but perhaps the degree to which policy makers are comfortable with the economic evaluation paradigm as a whole. Cost-effectiveness is indeed only one of several considerations that policy makers face in considering the adoption of a new intervention (Musgrove 1999). DSTs can however be developed using multi-criteria decision analysis frameworks, which account for a broader range of factors beyond cost-effectiveness, such as equity, acceptability, or the degree to which an intervention is life-saving (Dowie 2008).

In Chapter 7, the use of multilevel modelling, a relatively recent tool in economic evaluation, was demonstrated for the evaluation of a multi-centre trial. Multicentre trials are increasingly being used to ensure that trial results can claim greater generalisability. The interpretation of these results and the handling of variability of costs and effectiveness between locations have so far been inadequate, with either pooling or stratification of the data being the norm (Manca, Rice et al. 2005). The use of multilevel modelling offers a method that addresses the limitations of both approaches. While the analysis had a number of limitations, and could not undertake a full exploration of the benefits of MLM, it did provide an improved model when compared with the pooled and stratified models, as measured by the DIC.

In Chapter 8, the evaluation of the HMM programme was designed to explicitly show the variation in the intervention's efficiency across different sites. In addition to the variation in transmission intensity, as assessed in some previous evaluations (Goodman, Coleman et al. 2000; Shillcutt, Morel et al. 2008), this analysis explored the impact of varying levels of access to high quality care. The method used to portray differences in access to high quality care was crude, although it was adequate to show the significant impact that access to care has on the efficiency of the intervention, providing a clear visual demonstration of how results vary across different circumstances. The limitation of

presenting a broad array of results simultaneously, is the difficulty of carrying out extensive sensitivity analyses such as PSAs. This could perhaps be overcome with more advance use of Excel or the use of alternative software.

9.1.4 Appropriate model structures

Almost all evaluations based on modelling disease progression have structured their analyses using decision trees to model disease progression with the alternative treatments or diagnostics being considered (Cho Min, Lermaharit et al. 2000; Wilkins, Folb et al. 2002; Zurovac et al. 2006). Despite the strength and appealing simplicity of decision trees, their limitations in the context of malaria and its control are considerable. The linear portrayal of the illness does not easily facilitate the evaluation of recurrent events, and quickly becomes unwieldy where more than a couple of progression paths are present at each node.

Other evaluations have used mathematical modelling to assess disease progression and factor in changes in antimalarial effectiveness due to growing resistance (Schapira, Beals et al. 1993) and host susceptibility given changes in transmission and immunity (Laxminarayan 2004).

The Markov model presented in Chapter 8 is the first time such a model has been used in the economic evaluation of malaria treatment and diagnostics. Use of the model allowed for greater flexibility in the transmission of patients between health states than would be allowed by a decision tree structure. Other advantages of the model are the ability to assess the costs and benefits of the different strategies over extensive periods of time rather than those of a single event. Although in this instance the young age of the target population meant that modelling the development of immunity was not necessary, this is a dynamic that could be incorporated in the model with relative ease.

9.2 Sources of data

This thesis used different datasets for each analysis, rather than a single dataset throughout. This enabled the thesis to build models based on a range of policy relevant

issues already being explored, and to draw at least some of its data from recent or ongoing trials. A number of collaborations were established, through offering assistance with economic evaluation of existing trials in exchange for use of the data in the thesis. The advantage was that each study presented its own challenges in identifying appropriate modelling techniques. Furthermore, covering a range of interventions and strategies allowed for rapid adaptation of methods from one analysis into another.

There were two limitations to the use of data from multiple trials. First, not all trials had planned for an economic evaluation component, thus it had to be improvised after the trials had begun or even after their completion, such as in the evaluation of artesunate for the treatment of severe malaria. In some cases the data had to be collected retrospectively or estimated using costs of similar interventions and health facilities to those originally used in the trials. Second, there was less opportunity to explore both the methodological and practical issues of a single trial in greater depth, limiting the extent to which practical conclusions and policy guidelines could be produced.

9.3 Main policy findings and implications

While the orientation of the thesis is mostly methodological, a number of findings do have strong policy implications.

The incorporation of non-adherence to test results into the evaluation of RDTs had a decisive influence on their cost-effectiveness, in many instances severely curtailing it. This was particularly true for high transmission settings that characterise much of SSA where RDTs are being considered for use. The analysis identified levels of non-adherence at which RDTs cease to be efficient, rendering their use a misallocation of scarce resources. This strongly emphasises the need for clinician training programmes before widespread deployment of RDTs. Such programmes will of course have their own costs and benefits, which can be evaluated in light of the study developed here.

198

Similarly, the incorporation of the harm of treatment factor had a substantial impact on whether RDTs are justifiable on economic grounds. The explicit incorporation of an indirect cost associated with the use of antimalarials provided a necessary restraint on the use of antimalarials, a factor that had previously been addressed externally to the evaluation of diagnostics and antimalarials. Including this parameter formalises one of the main incentives of introducing RDTs – the reduction of unnecessary use of antimalarials. Its inclusion makes RDTs more beneficial than would otherwise be the case, for instance amongst children in low and medium transmission areas. By incorporating this parameter explicitly in the analysis, policy makers can directly observe its influence when considering the sub-populations in which RDTs are to be deployed.

The use of DSTs is likely to hold a number of advantages for policy makers over the traditional use of published evaluations. The main conclusion from the work on DSTs is their demonstration of the variability in results in response to the settings to which they are applied, as shown in Chapters 6 and 8. There rarely are simple policy recommendations that will hold true over such variable settings as those that characterise SSA. It is recognised that it would be impossible to tailor policies to highly localised levels. The level at which policy making can be independently made will vary according to factors such as training, drug supply systems and political constraints. Nevertheless, policy makers should at least be aware that blanket application of interventions might be inappropriate.

In Chapter 7 the focus was on the differences in expenditure of treating severe malaria with either artesunate or quinine. The supplementary study presented in Annex 3 found artesunate to be a highly cost-effective intervention, with a cost per death averted of approximately $140. While standard methods showed that artesunate would certainly be more expensive than quinine, when using MLM the difference was shown to be smaller and not statistically significant, partly alleviating the concern that despite being cost-effective, artesunate might not be affordable.

In Chapter 8 the Markov Model evaluated the efficiency of HMM programmes in a range of circumstances. Despite strong policy support for such programmes, the model results suggest that blanket application of this strategy will not always produce the best health

outcomes, nor will it always be economically efficient. This is likely to be particularly true given the current trend of reduced malaria prevalence, and ongoing urbanization improving access to care, in much of SSA.

9.4 Relevance to non-malarial contexts

While the focus of the analyses in the thesis were all malaria related interventions, most of the methods used are applicable to other contexts. The factors introduced into the evaluation of diagnostic tests – both non-adherence and the harm of treatment - were both influential and are highly relevant to the efficiency of diagnostic tests in many other fields.

DSTs offer a possible progression for economic evaluation in decentralizing health systems. In addition to ensuring the relevance of the evaluations to different locations and specific interventions, the use of DSTs facilitates re-visiting previous evaluations to verify that interventions continue to provide a justifiable return on the investment. This is particularly crucial in highly dynamic environments that characterise the developing world.

The use of multilevel models in the analysis of multi-centre trials has also been scarce in the context of economic evaluations, despite their demonstrated advantage over standard methods. With the increasing number of clinical trials being carried out in multiple locations, it is imperative that these use best available analytical practices as both pooling and stratifying results can lead to incorrect inferences.

9.5 Further research

A number of areas can be identified for further research, some relating to methodological advances, and others to the application of the methods used in this thesis to other aspects of malaria control.

Methodologically, much of the work in this thesis was aimed at increasing the comprehensiveness of analyses. Most of the costs and consequences in these analyses, however, were estimated at the micro level – patients, health care providers etc., without allowing for broader impacts, such as productivity and educational attainment. An

alternative to this is a macroeconomic approach which estimates the impact of both the disease and its control on the economy as a whole, including factors such as total productivity, economic growth, or the health sector as a whole (Chima, Goodman et al 2003).

The use of Computable General Equilibrium (CGE) methods is one possible method of estimating the costs and consequences of malaria related policies at the macro level, and has been demonstrated for use in the context of the development of antimicrobial resistance to drugs (Smith and Coast 2005) and of a global influenza pandemic (Keogh-Brown, McDonald et al. Unpublished). Policy analysis in CGE models functions by constructing a set of relationships between different sectors of the economy and the elasticities attached to these, and then assessing what would happen if some policy is introduced as an exogenous 'shock' to the model. The 'post-change' counterfactual equilibrium is then compared with the benchmark equilibrium to assess the costs and consequences of the policy in terms of macro-economic indicators – national income, employment, inflation or other welfare criteria.

Another area requiring further research is the use of decision models to determine the expected value of perfect information (EVPI) (Spiegelhalter, Abrams and Myles 2004). This involves estimating the impact of uncertainty surrounding a parameter on results, and using this to determine how much should be spent on reducing this uncertainty through further research. In the context of malaria this has particular relevance as many of the funding bodies face ongoing choices on whether to spend more on research or on deployment of the interventions. The use of decision models and EVPI can assist in determining this.

The use of multilevel modelling in this thesis served mostly to demonstrate its applicability to multi-centre trials. A similar trial to the present one in this analysis is currently under way in SSA, with MLM now planned as an integrated part of the analysis. Many of the limitations encountered here should be surmounted – with information collected prospectively on costs at both the individual and hospital level. The analysis will also account for effectiveness using a more advanced model than the one developed here.

An area receiving much recent attention is modelling the feasibility of regional and global malaria eradication programmes, and their supplementation with economic models to assess the efficiency of such programmes. The challenges involved are immense given the uncertainty surrounding the trajectories that malaria epidemiology could take in response to large scale elimination and eradication programmes. One of the main questions arising is whether a 'meta-model' for this is conceivable, or whether smaller models looking at different aspects of eradication should be left segregated.

While the challenges of drawing together a range of different models and analyses would be immense, this remains a venture that would be worth exploring as it could provide a better platform for predicting the results of new interventions and strategies than leaving the final convergence of the different inputs of models to be done in an intuitive manner. The linking of biological and economic models, such as that developed by Yeung (2006), show this is feasible, although further developments will be required to adapt such models to the context of malaria eradication.

Related to the evaluation of eradication programmes, is the need to evaluate different combinations of interventions rather than trying to distil the costs and benefits of specific ones. In practice the costs of interventions will often depend on whether they are deployed in combination with others as overhead costs may be shared, and the effectiveness of different interventions can be greater in combination than their sum total when implemented alone, or conversely they can cancel each other out. Either way these interactions need to be captured in decision models if the true costs and consequences in real life are to be estimated.

The analysis in Chapter 5 showed the decisive influence that adherence to test results had on efficiency. A trial currently under way in Uganda is assessing the effectiveness of a clinician training programme to encourage better adherence to negative test results, with promising interim outcomes (H. Hopkins, personal communication). The cost-effectiveness of such programmes can be assessed based on the degree to which they improve adherence and therefore the efficiency of RDTs.

Chapter 5 also explored the harm of treatment factor. As discussed, there is considerable uncertainty surrounding this parameter, mostly as a result of the difficulties in modelling the initial emergence of resistance to antimalarials, and its subsequent spread. This is an issue of critical importance for any modelling of the cost-effectiveness of malaria diagnostics and treatments, demanding greater attention.

9.6 Conclusion

Over three billion people across the globe are at risk of malaria (Snow, 2004). Estimates of malaria mortality range between 1 and 3 million people per year. At different places and times across much of the globe, malaria has been believed to be the cause of severe social and economic stagnation. The magnitude of the threat has prompted global efforts drawing at times on extensive resources, most notably the Global Malaria Eradication Programme. Although ultimately the GMEP was deemed a failure, eradication and control efforts in endemic areas have continued to draw in extensive proportions of national health budgets and international aid money for much of recent decades, and in the past few years this has reached an unprecedented scale, with strong support from bodies like the GFATM.

However, the influx of resources to control malaria, and the economic evaluations attempting to assist in directing these, have not always ensured best decision making practices with respect to the management of malaria and other febrile illnesses. For example, one consequence of the high profile malaria has attained is its overdiagnosis, and the mismanagement of other febrile illnesses such as pneumonia that can be extremely difficult to distinguish from malaria and carry higher mortality rates than malaria. Reflecting this tendency, economic evaluations of malaria interventions have often excluded health outcomes for patients with non-malarial illnesses who under strategies of presumptive treatment received antimalarials rather than treatment for their true cause of illness.

The high profile malaria control has attained in the aid sector and in the media has also encouraged the proclamation of magic bullet solutions, ignoring the variation in epidemiology, demography and environment that influence the efficiency of interventions. Such overly simplistic attitudes are also reflected in evaluations that have arguably over generalized their results beyond specific interventions and circumstances from which they drew their data.

The same sense of urgency to distribute new interventions has led to insufficient attention being paid to factors that are not immediately observed in standard trials, although they might have significant impacts on the true costs and benefits of the new interventions.

Control of malaria has evolved and expanded and the need for best available decision making tools has never been greater. The methods used to evaluate malaria related interventions, however, have not benefitted from recent developments in economic evaluation and decision analysis methods. This thesis has attempted to contribute to filling this gap, drawing on recent modelling approaches to improve the economic evaluation of the latest tools and strategies for the diagnosis and treatment of malaria.

References

Agnamey, P., P. Brasseur, et al. (2005). "Economic evaluation of a policy change from single-agent treatment for suspected malaria to artesunate-amodiaquine for microscopically confirmed uncomplicated falciparum malaria in the Oussouye District of south-western Senegal." Trop Med Int Health **10**(9): 926-33.

Adjuik, M., P. Agnamey, et al. (2002). "Amodiaquine-artesunate versus amodiaquine for uncomplicated Plasmodium falciparum malaria in African children: a randomised, multicentre trial." The Lancet **359**(9315): 1365-1372.

Ajayi, I. O., E. N. Browne, et al. (2008). "Feasibility and acceptability of artemisinin-based combination therapy for the home management of malaria in four African sites." Malar J **7**: 6.

Akhavan, D., P. Musgrove, et al. (1999). "Cost-effective malaria control in Brazil Cost-effectiveness of a Malaria Control Program in the Amazon Basin of Brazil, 1988–1996." Social Science & Medicine **49**(10): 1385-1399.

Al Serouri, A. W., S. M. Grantham-McGregor, et al. (2000). "Impact of asymptomatic malaria parasitaemia on cognitive function and school achievement of schoolchildren in the Yemen Republic." Parasitology **121 (Pt 4)**: 337-45.

Alilio, M. S., I. C. Bygbjerg, et al. (2004). "Are multilateral malaria research and control programs the most successful? Lessons from the past 100 years in Africa." Am J Trop Med Hyg **71**(2 Suppl): 268-78.

Alilio, M. S., A. Kitua, et al. (2004). "Malaria control at the district level in Africa: The case of the Muheza district in northeastern Tanzania." Am J Trop Med Hyg **71**(2): 205-213.

Alonzo Gonzalez, M., C. Menendez, et al. (2000). "Cost-effectiveness of iron supplementation and malaria chemoprophylaxis in the prevention of anaemia and malaria among Tanzanian infants." Bull World Health Organ **78**(1): 97-107.

Altman, D. G. and J. M. Bland (1994). "Diagnostic tests 2: Predictive values." BMJ **309**(6947): 102.

Amexo, M., R. Tolhurst, et al. (2004). "Malaria misdiagnosis: effects on the poor and vulnerable." Lancet **364**(9448): 1896-8.

Amin, A. A., D. A. Hughes, et al. (2004). "The difference between effectiveness and efficacy of antimalarial drugs in Kenya." Trop Med Int Health **9**(9): 967-74.

Anstey, N. M., R. N. Price, et al. (2006). "Improving the availability of artesunate for treatment of severe malaria." Med J Aust **184**(1): 3-4.

Appawu, M. A., K. M. Bosompem, et al. (2003). "Detection of malaria sporozoites by standard ELISA and VecTest™ dipstick assay in field-collected anopheline mosquitoes from a malaria endemic site in Ghana." Trop Med Int Health **8**(11): 1012-1017.

Arnesen, T. and L. Kapiriri (2004). "Can the value choices in DALYs influence global priority-setting?" Health Policy **70**(2): 137-49.

Arrow, J. K., C. B. Panosian, et al. (2004). Saving Lives, Buying Time: Economics of Malaria Drugs in an Age of Resistance.

Avila, P. E., K. Kirchgatter, et al. (2002). "Evaluation of a rapid dipstick test, Malar-Check™, for the diagnosis of *Plasmodium* falciparum malaria in Brazil." Revista do Instituto de Medicina Tropical de São Paulo **44**(5): 293-296.

Azazy, A. A. (2004). "Performance and accuracy of an immunodiagnostic antigen detection test in diagnosing*Plasmodium* falciparumamong Yemeni patients." Annals of Saudi Medicine **24**(1): 50-51.

Baird, J. K. (2005). "Effectiveness of antimalarial drugs." N Engl J Med 352(15): 1565-77.

Baltussen, R., W. Brouwer, et al. (2005). "Cost-effectiveness analysis for priority setting in health: penny-wise but pound-foolish." Int J Technol Assess Health Care **21**(4): 532-4.

Baltussen, R. and L. Niessen (2006). "Priority setting of health interventions: the need for multi-criteria decision analysis." Cost Eff Resour Alloc **4**: 14

Barat, L., J. Chipipa, et al. (1999). "Does the availability of blood slide microscopy for malaria at health centers improve the management of persons with fever in Zambia?" Am J Trop Med Hyg **60**(6): 1024-30.

Barker, C. and A. Green (1996). "Opening the debate on DALYs (disability-adjusted life years)." Health Policy Plan **11**(2): 179-83.

Barlow, R. (1968). "The economic effects of malaria eradication." American Economic Review **May**: 130-148.

Barnes, K. I., D. N. Durrheim, et al. (2005). "Effect of artemether-lumefantrine policy and improved vector control on malaria burden in KwaZulu-Natal, South Africa." PLoS Med **2**(11): e330.

Barnum, H. (1978). An Economic Analysis of a Malaria Control Program in the Outer Islands of Indonesia. Unpublished.

Barton, P., S. Bryan, et al. (2004). "Modelling in the economic evaluation of health care: selecting the appropriate approach." J Health Serv Res Policy 9(2): 110-8.

Bates, I., V. Bekoe, et al. (2004). "Improving the accuracy of malaria-related laboratory tests in Ghana." Malar J 3: 38.

Bates, N. and J. Herrington (2007). "Advocacy for malaria prevention, control, and research in the twenty-first century." Am J Trop Med Hyg 77(6 Suppl): 314-20.

Behrman, J. R., H. Alderman, et al., Eds. (2004). Hunger and Malnutrition: Copenhagen Consensus Challenge Paper. Global Crises Global Solutions. Cambridge, Cambridge University Press.

Bell, D. (2002). "Malaria rapid diagnostic tests: one size may not fit all." Clin Microbiol Rev 15(4): 771; discussion 771-2.

Bell, D. and R. W. Peeling (2006). "Evaluation of rapid diagnostic tests: malaria." Nat Rev Microbiol 4(9 Suppl): S34-8.

Bell, D. and M. D. Perkins (2008). "Making malaria testing relevant: beyond test purchase." Trans R Soc Trop Med Hyg.

Bell, D., C. Wongsrichanalai, et al. (2006). "Ensuring quality and access for malaria diagnosis: how can it be achieved?" Nat Rev Microbiol 4(9 Suppl): S7-20.

Bell, D. J. and M. E. Molyneux (2007). "Treatment of childhood Plasmodium falciparum malaria: current challenges." Expert Rev Anti Infect Ther 5(1): 141-52.

Bell, D. R., D. W. Wilson, et al. (2005). "False-positive results of a Plasmodium falciparum histidine-rich protein 2-detecting malaria rapid diagnostic test due to high sensitivity in a community with fluctuating low parasite density." Am J Trop Med Hyg 73(1): 199-203.

Berry, D. A. and D. K. Stangl (1996). Bayesian Biostatistics. New York, M. Dekker.

Bhattarai, A., A. S. Ali, et al. (2007). "Impact of Artemisinin-Based Combination Therapy and Insecticide-Treated Nets on Malaria Burden in Zanzibar." PLoS Med 4(11): e309.

Birch, S. and A. Gafni (1992). "Cost-effectiveness/utility analyses. Do current decision rules lead us to where we want to be?" J Health Econ 11(3): 279-96.

Biritwum, R. B., J. Welbeck, et al. (2000). "Incidence and management of malaria in two communities of different socio-economic level, in Accra, Ghana." Ann Trop Med Parasitol 94(8): 771-8.

Bloland, P. B., M. Ettling, et al. (2000). "Combination therapy for malaria in Africa: hype or hope?" Bull World Health Organ 78(12): 1378-88.

Bloland, P. B., S. P. Kachur, et al. (2003). "Trends in antimalarial drug deployment in sub-Saharan Africa." J Exp Biol 206(Pt 21): 3761-9.

Bobadilla, J. L., P. Cowley, et al. (1994). "Design, content and financing of an essential national package of health services." Bull World Health Organ 72(4): 653-62.

Bossert, T. (1998). "Analyzing the decentralization of health systems in developing countries: decision space, innovation and performance." Soc Sci Med 47(10): 1513-27.

Bottius, E., A. Guanzirolli, et al. (1996). "Malaria: even more chronic in nature than previously thought; evidence for subpatent parasitaemia detectable by the polymerase chain reaction." Trans R Soc Trop Med Hyg 90(1): 15-9.

Breman, J. G. (2001). "The ears of the hippopotamus: manifestations, determinants, and estimates of the malaria burden." Am J Trop Med Hyg 64(1-2 Suppl): 1-11.

Breman, J. G., M. S. Alilio, et al. (2004). "Conquering the intolerable burden of malaria: what's new, what's needed: a summary." Am J Trop Med Hyg 71(2 Suppl): 1-15.

Brennan, A. and R. Akehurst (2000). "Modelling in health economic evaluation. What is its place? What is its value?" Pharmacoeconomics 17(5): 445-59.

Brenner, H. and O. Gefeller (1997). "Variation of sensitivity, specificity, likelihood ratios and predictive values with disease prevalence." Stat Med 16(9): 981-91.

Brentlinger, P. E., C. B. Behrens, et al. (2006). "Challenges in the concurrent management of malaria and HIV in pregnancy in sub-Saharan Africa." Lancet Infect Dis 6(2): 100-11.

Briggs, A. (2000). "Economic evaluation and clinical trials: size matters." Bmj 321(7273): 1362-3.

Briggs, A. and M. Sculpher (1998). "An introduction to Markov modelling for economic evaluation." Pharmacoeconomics 13(4): 397-409.

Briggs, A., M. J. Sculpher, et al. (2006). Decision Modelling for Health Economic Evaluation, Oxford University Press.

Briggs, A. H. (2000). "Handling uncertainty in cost-effectiveness models." Pharmacoeconomics **17**(5): 479-500.

Brisson, M. and W. J. Edmunds (2006). "Impact of model, methodological, and parameter uncertainty in the economic analysis of vaccination programs." Med Decis Making **26**(5): 434-46.

Bruce-Chwatt, L. J. (1987). "Malaria and its control: present situation and future prospects." Annu Rev Public Health **8**: 75-110.

Bruce-Chwatt, L. J. (2002). Essential Malariology, Hodder Arnold.

Bruce, C. L. J. (1985). Essentials of malariology. New York, John Wiley and Sons.

Bryan, S. and J. Brown (1998). "Extrapolation of cost-effectiveness information to local settings." J Health Serv Res Policy **3**(2): 108-12.

Bryan, S., I. Williams, et al. (2007). "Seeing the NICE side of cost-effectiveness analysis: a qualitative investigation of the use of CEA in NICE technology appraisals." Health Econ **16**(2): 179-93.

Bualombai, P., S. Prajakwong, et al. (2003). "Determining cost-effectiveness and cost component of three malaria diagnostic models being used in remote non-microscope areas." The Southeast Asian J of Trop Med and Public Health **34**(2): 322-33.

Bustreo, F., A. Harding, et al. (2003). "Can developing countries achieve adequate improvements in child health outcomes without engaging the private sector?" World Health Organ **81**(12): 886-895.

Butterworth, T. and V. Bishop (1995). "Identifying the characteristics of optimum practice: findings from a survey of practice experts in nursing, midwifery and health visiting." J Adv Nurs **22**(1): 24-32.

Buxton, M. J., M. F. Drummond, et al. (1997). "Modelling in economic evaluation: an unavoidable fact of life." Health Econ **6**(3): 217-27.

Cancre, N., A. Tall, et al. (2000). "Bayesian analysis of an epidemiologic model of Plasmodium falciparum malaria infection in Ndiop, Senegal." Am J Epidemiol **152**(8): 760-70.

Ceesay, S. J., C. Casals-Pascual, et al. (2008). "Changes in malaria indices between 1999 and 2007 in The Gambia: a retrospective analysis." Lancet **372**(9649): 1545-54.
Centre for Multilevel Modelling (2008). LEMMA: Multilevel Modelling online course, CMM, University of Bristol.

Chanda, P., F. Masiye, et al. (2007). "A cost-effectiveness analysis of artemether lumefantrine for treatment of uncomplicated malaria in Zambia." Malar J 6: 21.

Chandler, C. I., C. Jones, et al. (2008). "Guidelines and mindlines: why do clinical staff over-diagnose malaria in Tanzania? A qualitative study." Malar J 7(1): 53.

Chandramohan, D., S. Jaffar, et al. (2002). "Use of clinical algorithms for diagnosing malaria." Trop Med Int Health 7(1): 45-52.

Chatburn, R. L. and F. P. Primiano, Jr. (2001). "Decision analysis for large capital purchases: how to buy a ventilator." Respir Care 46(10): 1038-53.

Chima, R. I., C. A. Goodman, et al. (2003). "The economic impact of malaria in Africa: a critical review of the evidence." Health Policy 63(1): 17-36.

Chiodini, P. L. (1998). "Non-microscopic methods for diagnosis of malaria." Lancet 351(9096): 80-1.

Chiodini, P. L., K. Bowers, et al. (2007). "The heat stability of *Plasmodium* lactate dehydrogenase-based and histidine-rich protein 2-based malaria rapid diagnostic tests." Trans R Soc Trop Med Hyg 101(4): 331-7.

Cho Min, N., S. Lertmaharit, et al. (2000). "Ex post and ex ante willingness to pay (WTP) for the ICT Malaria Pf/Pv test kit in Myanmar." Southeast Asian J Trop Med Public Health 31(1): 104-11.

Cho Min, N. and A. Saul (2000). "Treatment of uncomplicated *Plasmodium* falciparum malaria in Myanmar: a clinical decision analysis." Southeast Asian J Trop Med Public Health 31(2): 238-45.

Claxton, K., M. Sculpher, et al. (2005). "Probabilistic sensitivity analysis for NICE technology assessment: not an optional extra." Health Econ 14(4): 339-47

CMM. (2008). "LEMMA: Multilevel Modelling online course." Retrieved July 2008, 2008.

Coast, J. and R. D. Smith (2003). "Antimicrobial resistance: cost and containment." Expert Rev Anti Infect Ther 1(2): 241-51.

Cohn, E. (1973). "Assessing the costs and benefits of anti-malaria programs: The Indian experience." American J of Public Health 63(12): 1086-96.

Coleman, P. G., C. A. Goodman, et al. (1999). "Rebound mortality and the cost-effectiveness of malaria control: potential impact of increased mortality in late childhood following the introduction of insecticide treated nets." Trop Med Int Health 4(3): 175-86.

Coleman, P. G., C. Morel, et al. (2004). "A threshold analysis of the cost-effectiveness of artemisinin-based combination therapies in sub-saharan Africa." Am J Trop Med Hyg 71(2 Suppl): 196-204.

Cooper, N. J., A. J. Sutton, et al. (2004). "Comprehensive decision analytical modelling in economic evaluation: a Bayesian approach." Health Econ 13(3): 203-26.

Cooper, N. J., A. J. Sutton, et al. (2007). "Use of evidence in economic decision models: practical issues and methodological challenges." Health Econ 16(12): 1277-1286.

Cross, A. P. and B. Singer (1991). "Modelling the development of resistance of Plasmodium falciparum to anti-malarial drugs." Trans R Soc Trop Med Hyg 85(3): 349-55.

Cruciani, M., S. Nardi, et al. (2004). "Systematic review of the accuracy of the ParaSight-F test in the diagnosis of Plasmodium falciparum malaria." Med Sci Monit 10(7): MT81-8.

Curtis, C. F., B. Jana-Kara, et al. (2003). "Insecticide treated nets: impact on vector populations and relevance of initial intensity of transmission and pyrethroid resistance." J Vector Borne Dis 40(1-2): 1-8

D'Alessandro, U. (1999). "A rational approach to malaria control in pregnancy in sub-Saharan Africa: the need for a link between scientific research and public-health interventions." Ann Trop Med Parasitol 93 Suppl 1: S75-7.

D'Alessandro, U. (2001). "Malaria control in Tanzania." Lancet 358(9283): 762-3.

D'Alessandro, U. (2004). "Treating severe and complicated malaria." BMJ 328(7432): 155.

D'Alessandro, U. and H. Buttiens (2001). "History and importance of antimalarial drug resistance." Trop Med Int Health 6(11): 845-8.

D'Alessandro, U., A. Talisuna, et al. (2005). "Editorial: Should artemisinin-based combination treatment be used in the home-based management of malaria?" Trop Med Int Health 10(1): 1-2.

Daga, S. R., V. B. Patil, et al. (2005). "Syndromic management of prolonged fever: a cost-effective approach." Trop Doct 35(1): 31-4.

Davies, H. T. O. and S. M. Nutley (1999). "The Rise and Rise of Evidence in Health Care." Public Money and Management 19(1): 9-16.

Davis, T. M. E. (2000). "Recognition and management of falciparum malaria." Emergency Medicine 12(4): 276-284.

De Castro, M. C., Y. Yamagata, et al. (2004). "Integrated urban malaria control: A case study in Dar Es Salaam, Tanzania." Am J Trop Med Hyg **71**(2): 103-117.

de Savigny, D., C. Mayombana, et al. (2004). "Care-seeking patterns for fatal malaria in Tanzania." Malar J **3**: art. no.-27.

Dondorp, A., F. Nosten, et al. (2005). "Artesunate versus quinine for treatment of severe falciparum malaria: a randomised trial." Lancet **366**(9487): 717-25.

Dorsey, G., D. Njama, et al. (2002). "Sulfadoxine/pyrimethamine alone or with amodiaquine or artesunate for treatment of uncomplicated malaria: a longitudinal randomised trial." Lancet **360**(9350): 2031-8.

Dowie, J. (2003). Health impact: its estimation, assessment and analysis. Public health for the 21st century: new perspectives on policy, participation and practice. J. Orme, J. Powell, P. Taylor, T. Harrison and M. Gray, Open University Press.

Dowie, J. (2005). "Research Implications of Science-Informed, Value-Based Decision Making." Human and Ecological Risk Assessment, **11**(1).

Dowie, J. (2008) "The future of HTA is MCDA." http://knol.google.com/k/jack-dowie/the-future-of-hta-is-mcda/2rrcx2o0qipsa/14# Accessed 20/11/08

Drakeley, C. J., I. Carneiro, et al. (2005). "Altitude-dependent and independent variations in *Plasmodium* falciparum prevalence in northeastern Tanzania." Journal Infectious Diseases **191**(10): 1589-1598.

Drakeley, C. J., P. H. Corran, et al. (2005). "Estimating medium- and long-term trends in malaria transmission by using serological markers of malaria exposure." Proc Natl Acad Sci U S A **102**(14): 5108-13.

Drummond, M., G. Torrance, et al. (1993). "Cost-effectiveness league tables: more harm than good?" Soc Sci Med **37**(1): 33-40.

Drummond, M. F. and T. O. Jefferson (1996). "Guidelines for authors and peer reviewers of economic submissions to the BMJ. The BMJ Economic Evaluation Working Party." Bmj **313**(7052): 275-83.

Drummond, M. F., B. O'Brien, et al. (2005). Methods for the Economic Evaluation of Health Care Programmes, Oxford University Press.

Dunlop, D. W. (1984). "Theoretical and empirical issues in benefit identification, measurement and valuation related to parasitic disease control in poor countries." Soc Sci Med **19**(10): 1031-7.

Edejer, T. T., R. Baltussen, et al. (2003). WHO Guide to Cost-effectiveness Analysis.

English, M., C. Waruiru, et al. (1996). "Transfusion for respiratory distress in life-threatening childhood malaria." Am J Trop Med Hyg 55(5): 525-30.

English, M. and J. A. Scott (2008). "What Is the Future for Global Case Management Guidelines for Common Childhood Diseases?" PLoS Med 5(12)

Evans, D. (2003). "Hierarchy of evidence: a framework for ranking evidence evaluating health care interventions." J Clin Nurs 12(1): 77-84.

Evans, D. B., Ed. (2004). Perspective Paper 2.1. . Global Crises Global Solutions. Cambridge, Cambridge University Press.

Fapohunda, Bolaji M., Beth Ann Plowman, Robert Azairwe, Geoffrey Bisorbowa, Peter Langi,
Frederick Kato and Xiaotian Wang (2004). Home-Based Management of Fever Strategy in Uganda: A Report of the 2003 Survey. Arlington, Virginia, USA: MOH, WHO and BASICS II

Fenwick, E., K. Claxton, et al. (2000). Improving the efficiency and relevance of health technology assessment: the role of iterative decision analytic modelling, Centre for Health Economics: University of York.

Fenwick, E., B. J. O'Brien, et al. (2004). "Cost-effectiveness acceptability curves--facts, fallacies and frequently asked questions." Health Econ 13(5): 405-15.

Fernando, S. D., N. D. Karunaweera, et al. (2004). "A cost analysis of the use of the rapid, whole-blood, immunochromatographic P.f/P.v assay for the diagnosis of *Plasmodium* vivax malaria in a rural area of Sri Lanka." Ann Trop Med Parasitol 98(1): 5-13.

Foster, S. and M. Phillips (1998). "Economics and its contribution to the fight against Malaria." Ann Trop Med Parasitol 92(4): 391-398.

Francis, D., S. L. Nsobya, et al. (2006). "Geographic differences in antimalarial drug efficacy in Uganda are explained by differences in endemicity and not by known molecular markers of drug resistance." J Infect Dis 193(7): 978-86.

Frew, E. J., D. K. Whynes, et al. (2003). "Eliciting willingness to pay: comparing closed-ended with open-ended and payment scale formats." Med Decis Making 23(2): 150-9.

Fryback, D. G., J. O. Chinnis, Jr., et al. (2001). "Bayesian cost-effectiveness analysis. An example using the GUSTO trial." Int J Technol Assess Health Care 17(1): 83-97.

Fryback, D. G., N. K. Stout, et al. (2001). "An elementary introduction to Bayesian computing using WinBUGS." Int J Technol Assess Health Care 17(1): 98-113.

Gafni, A. and S. Birch (2006). "Incremental cost-effectiveness ratios (ICERs): the silence of the lambda." Soc Sci Med 62(9): 2091-100.

Gallup, J. L. and J. D. Sachs (2001). "The economic burden of malaria." Am J Trop Med Hyg 64(1-2 Suppl): 85-96.

Garber, A. M. and C. E. Phelps (1997). "Economic foundations of cost-effectiveness analysis." J Health Econ 16(1): 1-31.

Gerard, K. and G. Mooney (1993). "QALY league tables: handle with care." Health Econ 2(1): 59-64.

Gilles, H. M. and A. O. Lucas (1998). "Tropical medicine: 100 years of progress." Br Med Bull 54(2): 269-80.

Girosi, F., M. E. Rafael, et al. (2006). Determining the Priority Global Health Needs and Quantifying the Health Benefits Resulting From the Introduction of New Diagnostics in the Developing World. RAND Health. RAND.

Gogtay, N. J., V. S. Kadam, et al. (2003). "A cost-effectiveness analysis of three antimalarial treatments for acute, uncomplicated Plasmodium falciparum malaria in Mumbai, India." J Assoc Physicians India 51: 877-9.

Gold, M. R., J. E. Siegel, et al. (1996). Cost-effectiveness in Health and Medicine. New York, Oxford University Press.

Gold, M. R., D. Stevenson, et al. (2002). "HALYS and QALYS and DALYS, Oh My: similarities and differences in summary measures of population Health." Annu Rev Public Health 23: 115-34.

Goodman, C., P. G. Coleman, et al. (2000). Economic Analysis of Malaria Control in Sub-Saharan Africa. Strategic Research Series, Global Forum for Health Research.

Goodman, C., S. P. Kachur, et al. (2004). "Retail supply of malaria-related drugs in rural Tanzania: risks and opportunities." Trop Med Int Health 9(6): 655-63.

Goodman, C., W. Mutemi, et al. (2006). "The cost-effectiveness of improving malaria home management: shopkeeper training in rural Kenya." Health Policy Plan.

Goodman, C. A., P. G. Coleman, et al. (1999). "Cost-effectiveness of malaria control in sub-Saharan Africa." Lancet 354(9176): 378-85.

Goodman, C. A., P. G. Coleman, et al. (2001a). "Changing the first line drug for malaria treatment--cost-effectiveness analysis with highly uncertain inter-temporal trade-offs." Health Econ **10**(8): 731-49.

Goodman, C. A., P. G. Coleman, et al. (2001b). "The cost-effectiveness of antenatal malaria prevention in sub-Saharan Africa." Am J Trop Med Hyg **64**(1-2 Suppl): 45-56.

Goodman, C. A. and A. J. Mills (1999). "The evidence base on the cost-effectiveness of malaria control measures in Africa." Health Policy Plan **14**(4): 301-12.

Goodman, C. A., A. E. Mnzava, et al. (2001). "Comparison of the cost and cost-effectiveness of insecticide-treated bednets and residual house-spraying in KwaZulu-Natal, South Africa." Trop Med Int Health **6**(4): 280-95.

Goodman, C. A., W. M. Mutemi, et al. (2006). "The cost-effectiveness of improving malaria home management: shopkeeper training in rural Kenya." Health Policy Plan **21**(4): 275-88.

Gosling, R. D., C. J. Drakeley, et al. (2008). "Presumptive treatment of fever cases as malaria: help or hindrance for malaria control?" Malar J **7**: 132.

Greenwood, B. M., K. Bojang, et al. (2005). "Malaria." Lancet **365**(9469): 1487-98.

Grieve, R., R. Nixon, et al. (2007). "Multilevel models for estimating incremental net benefits in multinational studies." Health Econ **16**(8): 815-26.

Grieve, R., R. Nixon, et al. (2005). "Using multilevel models for assessing the variability of multinational resource use and cost data." Health Econ **14**(2): 185-96.

Griffith, D. H. S., D. V. Ramara, et al. (1971). "Contribution of Health to Development." International Journal of Health Services **1**(253-318).

Gupta, S., R. W. Snow, et al. (1999). "Immunity to non-cerebral severe malaria is acquired after one or two infections." Nat Med **5**(3): 340-3.

Guthmann, J. P., A. Ruiz, et al. (2002). "Validity, reliability and ease of use in the field of five rapid tests for the diagnosis of *Plasmodium* falciparum malaria in Uganda." Transactions Of The Royal Society Of Tropical Medicine And Hygiene **96**(3): 254-257.

Hamer, D. H., M. Ndhlovu, et al. (2007). "Improved diagnostic testing and malaria treatment practices in Zambia." JAMA **297**(20): 2227-31.

Hammer, J. S. (2001). "THE ECONOMICS OF MALARIA CONTROL." The World Bank Research Observer **8**(1): 1-22.

Hammond, K. R. (2000). Human Judgment and Social Policy: Irreducible Uncertainty, Inevitable Error. New York, Oxford University Press.

Hammond, K. R., R. M. Hamm, et al. (1997). 4 Direct comparison of the efficacy of intuitive and analytical cognition in expert judgment. Research on Judgment and Decision Making. W. M. Goldstein and R. M. Hogarth, Cambridge University Press

Hanscheid, T. (2003). "Current strategies to avoid misdiagnosis of malaria." Clinical Microbiology And Infection 9(6): 497-504.

Harvey, S. A., L. Jennings, et al. (2008). "Improving community health worker use of malaria rapid diagnostic tests in Zambia: package instructions, job aid and job aid-plus-training." Malar J 7: 160.

Hastings, I. M. (2003). "Malaria control and the evolution of drug resistance: an intriguing link." Trends Parasitol 19(2): 70-3.

Haynes, R. B. (2002). "What kind of evidence is it that Evidence-Based Medicine advocates want health care providers and consumers to pay attention to?" BMC Health Serv Res 2(1): 3.

Hetzel, M. W., N. Iteba, et al. (2007). "Understanding and improving access to prompt and effective malaria treatment and care in rural Tanzania: the ACCESS Programme." Malar J 6: 83.

Hien, T. T., N. P. J. Day, et al. (1996). "A Controlled Trial of Artemether or Quinine in Vietnamese Adults with Severe Falciparum Malaria." New England Journal of Medicine 335(2): 76.

Hoch, J. S., A. H. Briggs, et al. (2002). "Something old, something new, something borrowed, something blue: a framework for the marriage of health econometrics and cost-effectiveness analysis." Health Econ 11(5): 415-30.

Honrado, E. R., W. Fungladda, et al. (1999). "Cost-effectiveness analysis of artesunate and quinine + tetracycline for the treatment of uncomplicated falciparum malaria in Chanthaburi, Thailand." Bull World Health Organ 77(3): 235-43.

Hopkins, H., L. Bebell, et al. (2008). "Rapid diagnostic tests for malaria at sites of varying transmission intensity in Uganda." J Infect Dis 197(4): 510-8.

Hopkins, H., W. Kambale, et al. (2007). "Comparison of HRP2- and pLDH-based rapid diagnostic tests for malaria with longitudinal follow-up in Kampala, Uganda." Am J Trop Med Hyg 76(6): 1092-1097.

Hopkins, H., A. Talisuna, et al. (2007). "Impact of home-based management of malaria on health outcomes in Africa: a systematic review of the evidence." Malar J 6: 134.

Hoshen, M. (2004). "Artesunate combinations for malaria." Lancet 363(9410): 737.

Hutubessy, R. C., L. M. Bendib, et al. (2001). "Critical issues in the economic evaluation of interventions against communicable diseases." Acta Trop 78(3): 191-206.

Iqbal, J., A. Muneer, et al. (2003). "Performance of the optimal test for malaria diagnosis among suspected malaria patients at the rural health centers." Am J Trop Med Hyg 68(5): 624-628.

James, C., G. Carrin, et al. (2005). "Clarifying Efficiency-Equity Tradeoffs Through Explicit Criteria, With a Focus on Developing Countries." Health Care Analysis 13(1).

Jamison, D. T., W. H. Mosely, et al. (1993). Disease control priorities in developoing countries. Oxford, Oxford University Press.

Jha, P., A. Mills, et al. (2002). "Improving the health of the global poor." Science 295(5562): 2036-9.
Johann-Liang, R. and R. Albrecht (2003). "Safety evaluations of drugs containing artemisinin derivatives for the treatment of malaria." Clin Infect Dis 36(12): 1626-7; author reply 1627-8.

Jones, K. L., S. Donegan, et al. (2007). "Artesunate versus quinine for treating severe malaria." Cochrane Database Syst Rev(4): CD005967.

Jonkman, A., R. A. Chibwe, et al. (1995). "Cost-saving through microscopy-based versus presumptive diagnosis of malaria in adult outpatients in Malawi." Bull World Health Organ 73(2): 223-7.

Jowett, M. and N. J. Miller (2005). "The financial burden of malaria in Tanzania: implications for future government policy." Int J Health Plann Manage 20(1): 67-84.

Kachur, S. P., R. A. Khatib, et al. (2004). "Adherence to antimalarial combination therapy with sulfadoxine-pyrimethamine and artesunate in rural Tanzania

Retail supply of malaria-related drugs in rural Tanzania: risks and opportunities

Trends in antimalarial drug deployment in sub-Saharan Africa." Am J Trop Med Hyg 71(6): 715-22.

Kaewsonthi, S., A. G. Harding, et al. (1996). "Assessing the economic impact of a rapid on-site malaria diagnostic test." Southeast Asian J Trop Med Public Health 27(2): 210-5.

Kager, P. A. (2002). "Malaria control: constraints and opportunities." Trop Med Int Health 7(12): 1042-6.

Kallander, K., H. Hildenwall, et al. (2008). "Delayed care seeking for fatal pneumonia in children aged under five years in Uganda: a case-series study." Bull World Health Organ 86(5): 332-8.
Kassirer, J. P. and M. Angell (1994). "The journal's policy on cost-effectiveness analyses." N Engl J Med 331(10): 669-70.

Keogh-Brown, M., S. McDonald, et al. (Unpublished).
"The macroeconomic costs of a global influenza pandemic "

Kidane, G. and R. H. Morrow (2000). "Teaching mothers to provide home treatment of malaria in Tigray, Ethiopia: a randomised trial." Lancet 356(9229): 550-5.

Kihara, M., J. A. Carter, et al. (2006). "The effect of Plasmodium falciparum on cognition: a systematic review." Trop Med Int Health 11(4): 386-97.

Kindermans, J. (2004). Changing national malaria treatment protocols in Africa: What is the cost and who will pay? . Campaign for Access to Essential Medicines. MSF.

Kirkwood, B. R. and J. A. C. Sterne (2003). Essential Medical Statistics, Blackwell Publishing.

Kublin, J. G. and R. W. Steketee (2006). "HIV infection and malaria--understanding the interactions." J Infect Dis 193(1): 1-3.

Labelle, R. J. and J. E. Hurley (1992). "Implications of basing health-care resource allocations on cost-utility analysis in the presence of externalities." J Health Econ 11(3): 259-77.

Larson, B. A., A. A. Amin, et al. (2006). "The cost of uncomplicated childhood fevers to Kenyan households: implications for reaching international access targets." BMC Public Health 6: 314.

Lauer, J. A., K. Rohrich, et al. (2003). "PopMod: a longitudinal population model with two interacting disease states." Cost Eff Resour Alloc 1(1): 6.

Laufer, M. K., J. J. van Oosterhout, et al. (2006). "Impact of HIV-associated immunosuppression on malaria infection and disease in Malawi." J Infect Dis 193(6): 872-8.

Laxminarayan, R. (2004). "Act now or later? Economics of malaria resistance." Am J Trop Med Hyg 71(2 Suppl): 187-95.

Laxminarayan, R. and M. L. Weitzman (2002). "On the implications of endogenous resistance to medications." J Health Econ 21(4): 709-18.

Liberatore, M. J. and R. L. Nydick (In press). "The analytic hierarchy process in medical and health care decision making: A literature review." European Journal of Operational Research.

Lindley, D. V. (1985). Making Decisions. London, Wiley.

Lines J, Schapira A, Smith T. (2008) "Tackling malaria today." Bmj **337a**(869):435-437

Lindley, D. V. (2000). "The Philosophy of Statistics." The Statistician **49**(3): 293-337.

Lubell, Y., H. Reyburn, et al. (2007a). "The cost-effectiveness of parasitological diagnosis for malaria-suspected patients in an era of combination therapy." Am J Trop Med Hyg **77**(Suppl 6): 128-132.

Lubell, Y., H. Hopkins, et al. (2007b). Modelling Costs and Benefits of RDTs for the Detection of Plasmodium falciparum in Uganda. American Society of Tropical Medicine and Hygiene 56th Annual Meeting. Philadelphia.

Lubell, Y., H. Reyburn, et al. (2008a). "The impact of response to the results of diagnostic tests for malaria: cost-benefit analysis." Bmj **336**(7637): 202-5.

Lubell, Y., H. Hopkins, et al. (2008b). "An interactive model for the assessment of the economic costs and benefits of different rapid diagnostic tests for malaria." Malar J **7**: 21.

Lubell, Y., Yeung S., Dondorp A. M., et al. (In press). Cost-effectiveness of artesunate for the treatment of severe malaria. Trop Med Int Health.

Lucas, A. M. (2006). "Economic Effects of Malaria Eradication: Evidence from the Malarial Periphery." Unpublished Manuscript.

Lyttkens, C. H. (2003). "Time to disable DALYs? On the use of disability-adjusted life-years in health policy." Eur J Health Econ **4**(3): 195-202.

Macauley, C. (2005). "Aggressive active case detection: a malaria control strategy based on the Brazilian model." Soc Sci Med **60**(3): 563-73.

Makani, J., W. Matuja, et al. (2003). "Admission diagnosis of cerebral malaria in adults in an endemic area of Tanzania: implications and clinical description." Qjm **96**(5): 355-62.

Malaney, P., A. Spielman, et al. (2004). "The malaria gap." Am J Trop Med Hyg **71**(2 Suppl): 141-6.

Malaria Consortium (2003). Desk review for RBM essential actions progress... Investment gaps. USAID.

Malik, E. M., T. A. Mohamed, et al. (2006). "From chloroquine to artemisinin-based combination therapy: the Sudanese experience." Malar J 5: 65.

Malik, S., S. Khan, et al. (2004). "*Plasmodium* lactate dehydrogenase assay to detect malarial parasites." Natl Med J India 17(5): 237-9.

Manca, A., N. Rice, et al. (2005). "Assessing generalisability by location in trial-based cost-effectiveness analysis: the use of multilevel models." Health Econ 14(5): 471-85.

Marseille, E., J. Saba, et al. (2006). "The costs and benefits of private sector provision of treatment to HIV-infected employees in Kampala, Uganda." Aids 20(6): 907-14.

Marsh, K. and R. W. Snow (1999). "Malaria transmission and morbidity." Parassitologia 41(1-3): 241-6.

Mathers, C. D., C. Bernard, et al. (2002). Global Burden of Disease in 2002: data sources, methods and results. Global Programme on Evidence for Health Policy Discussion Paper No. 54. WHO.

Mayxay, M., S. Pukrittayakamee, et al. (2001). "Persistence of *Plasmodium* falciparum HRP-2 in successfully treated acute falciparum malaria." Trans Royal Soc Trop Med Hyg 95(2): 179-182.

Mayxay, M. F., P. N. Newton, et al. (2004). "An assessment of the use of malaria rapid tests by village health volunteers in rural Laos." Trop Med Int Health 9(3): 325-329.

Mbogo, C. N., R. W. Snow, et al. (1993). "Low-level *Plasmodium* falciparum transmission and the incidence of severe malaria infections on the Kenyan coast." Am J Trop Med Hyg 49(2): 245-53.

McIntosh, E., C. Donaldson, et al. (1999). "Recent advances in the methods of cost-benefit analysis in health care. Matching the art to the science." Pharmacoeconomics 15(4): 357-67.

McMarthy, F. D., H. Wolf, et al. (2000). Malaria and Growth. Policy Research Working Paper 2303, The World Bank.

McPake, B., L. Kumaranayake, et al. (2002). Health Economics - An International Perspective. London and New York, Routledge.

Mills, A. (1989). THE APPLICATION OF COST-EFFECTIVENESS ANALYSIS TO DISEASE CONTROL PROGRAMMES IN DEVELOPING COUNTRIES, WITH SPEICAL REFERENCE TO MALARIA CONTROL IN NEPAL. Evaluation and Planning Centre. London, London School of Hygiene and Tropical Medicine. PhD Thesis.

Mills, A. (1991). The economics of malaria control. Malaria: waiting for the vaccine. G. A. T. Targett. Chichester and New York, John Wiley and Sons: 141-68.

Mills, A. and S. D. Shillcutt, Eds. (2004). Communicable Diseases: Copenhagen Consensus Challenge Paper. Global Crises Global Solutions. Cambridge, Cambridge University Press.

MoH Tanzania (2002). National Malaria Medium Strategic Plan 2002-2007. Dar-es-Salaam.

Molineaux, L. and Gramiccia, G. (1980) The Garki project research on the epidemiology of malaria in the Sudan Savanna of West Africa. WHO, Geneva

Molineaux, L., D. A. Muir, et al. (1988). The epidemiology of malaria and its measurement. Malaria, principles and practice of Malariology. Edinburgh, Churchill Livingstone: 999-1090.

Mont, D. (2007). "Measuring health and disability." Lancet **369**(9573): 1658-63.

Moody, A. (2002). "Rapid diagnostic tests for malaria parasites." Clin Microbiol Rev **15**(1): 66-78.

Moody, A. H. and P. L. Chiodini (2002). "Non-microscopic method for malaria diagnosis using OptiMAL IT, a second-generation dipstick for malaria pLDH antigen detection." British Journal of Biomedical Science **59**(4): 228-231.

Moorthy, V. S., M. F. Good, et al. (2004). "Malaria vaccine developments." Lancet **363**(9403): 150-6.

Morel, C. M., J. A. Lauer, et al. (2005). "Cost-effectiveness analysis of strategies to combat malaria in developing countries." BMJ.

Morris, S., N. Devlin, et al. (2007). Economic Analysis in Health Care. Wiley and Sons Ltd.

Mortimer, D., R. D. Smith, et al. (2003). Development of an Economic Model of Antimicrobial Resistance, University of East Anglia.

MSH. (2007). International Drug Price Indicator Guide, http://erc.msh.org/dmpguide.

Mubyazi, G. M. and M. A. Gonzalez-Block (2005). "Research influence on antimalarial drug policy change in Tanzania: case study of replacing chloroquine with sulfadoxine-pyrimethamine as the first-line drug." Malar J **4**: 51.

Mugittu, K., M. Ndejembi, et al. (2004). "Therapeutic efficacy of sulfadoxine-pyrimethamine and prevalence of resistance markers in tanzania prior to revision of malaria treatment policy: *Plasmodium* falciparum dihydrofolate reductase and dihydropteroate synthase mutations in monitoring in vivo resistance." Am J Trop Med Hyg **71**(6): 696-702.

Muheki, C., D. McIntyre, et al. (2004). "Artemisinin-based combination therapy reduces expenditure on malaria treatment in KwaZulu Natal, South Africa." Trop Med Int Health 9(9): 959-66.

Mulligan, J., C. Morel, et al. (2005). The Cost-Effectiveness of Malaria Control Interventions. DISEASE CONTROL PRIORITIES PROJECT.

Mumba, M., J. Visschedijk, et al. (2003). "A Piot model to analyse case management in malaria control programmes." Trop Med Int Health 8(6): 544-551.

Murray, C. J., D. B. Evans, et al. (2000). "Development of WHO guidelines on generalized cost-effectiveness analysis." Health Econ 9(3): 235-51.

Murray, C. J. L. (2003). WHO Guide to Cost-Effectiveness Analysis. Geneva.

Murray, C. J. L. and A. Lopez (1996). The Global Burden of Disease: A comprehensive assessment of mortality and disability from diseases, injuries and risk factors in 1990 and projected to 2020, Harvard University Press.

Murray, C. K., D. Bell, et al. (2003). "Rapid diagnostic testing for malaria." Trop Med Int Health 8(10): 876-883.

Murray CK, Gasser RA, Jr., Magill AJ, Miller RS. (2008)"Update on rapid diagnostic testing for malaria." Clin Microbiol Rev;21(1):97-110.

Musgrove, P. (1999). "Public spending on health care: how are different criteria related?" Health Policy 47(3): 207-23.
Mutabingwa, T. K. (2001). "Monitoring antimalarial drug resistance within National Malaria Control Programmes: the EANMAT experience." Trop Med Int Health 6(11): 891-898.

Mutabingwa, T. K., D. Anthony, et al. (2005). "Amodiaquine alone, amodiaquine+sulfadoxine-pyrimethamine, amodiaquine+artesunate, and artemether-lumefantrine for outpatient treatment of malaria in Tanzanian children: a four-arm randomised effectiveness trial." Lancet 365(9469): 1474-80.

Mwakibinga, H., M. Mbvundula, et al. (2003). The profile and productivity of The National Malaria Control Programme. M. Tanzania.

Mwangi, T. W., A. Ross, et al. (2005). "Case definitions of clinical malaria under different transmission conditions in Kilifi District, Kenya." J Infect Dis 191(11): 1932-9.

Ndyomugyenyi, R., P. Magnussen, et al. (2007). "Diagnosis and treatment of malaria in peripheral health facilities in Uganda: findings from an area of low transmission in south-western Uganda." Malar J 6: 39.

Nixon, R. M. and S. G. Thompson (2004). "Parametric modelling of cost data in medical studies." Stat Med **23**(8): 1311-31.

NMCP (2005). Diagnosis and Treatment of Malaria Orientation Guide for District Trainers Malaria control series 10. U. R. o. T. Ministry of Health.

Nosten, F. and E. Ashley (2004). "The detection and treatment of *Plasmodium* falciparum malaria: Time for change." J Postgrad Med **50**(1): 35-39.

Novartis (2006). Novartis announces initiative to improve access to state-of-the-art anti-malarial treatment Coartem® Media Release. Basel.

O'Hagan, A., J. W. Stevens, et al. (2000). "Inference for the cost-effectiveness acceptability curve and cost-effectiveness ratio." Pharmacoeconomics **17**(4): 339-49.

Obua, C., L. L. Gustafsson, et al. (2006). "Improved efficacy with amodiaquine instead of chloroquine in sulfadoxine/pyrimethamine combination treatment of falciparum malaria in Uganda: experience with fixed-dose formulation." Acta Trop 100(1-2): 142-50.

Obua, C., L. L. Okell, L. C., C. J. Drakeley, et al. (2008). "Reduction of transmission from malaria patients by artemisinin combination therapies: a pooled analysis of six randomized trials." Malar J **7**: 125.

Okello, D. O. (1994). "Resource utilization patterns in patients with acquired immunodeficiency syndrome (AIDS)." East Afr Med J **71**(12): 816-7.

Olivar, M., M. Develoux, et al. (1991). "Presumptive diagnosis of malaria results in a significant risk of mistreatment of children in urban Sahel." Trans R Soc Trop Med Hyg **85**(6): 729-30.

Onwujekwe, O., R. Chima, et al. (2001). "Hypothetical and actual willingness to pay for insecticide-treated nets in five Nigerian communities." Trop Med Int Health **6**(7): 545-53.

Onwujekwe, O., B. Uzochukwu, et al. (2004). "Is combination therapy for malaria based on user-fees worthwhile and equitable to consumers? Assessment of costs and willingness to pay in Southeast Nigeria." Acta Trop **91**(2): 101-15.

Pandya, A. P., G. C. Sahu, et al. (2001). "The Para Check-PC Test: - a simple rapid dip stick test to detect <i>Plasmodium falciparum</i> infection." Journal of Communicable Diseases **33**(3): 224-225.

Panosian, C. B. (2005). "Economic access to effective drugs for falciparum malaria." Clin Infect Dis **40**(5): 713-7.

Pauly, M. V. (1993). Valuing health care benefits in money terms. Valuing Health Care. F. Sloan, Cambridge University Press: 99-124.

Phelps, C. E. and A. I. Mushlin (1991). "On the (near) equivalence of cost-effectiveness and cost-benefit analyses." Int J Technol Assess Health Care 7: 12.

Phillips, M. and P. A. Phillips-Howard (1996). "Economic implications of resistance to antimalarial drugs." Pharmacoeconomics 10(3): 225-38.

Picard, J., A. Mills, et al. (1992). "The cost-effectiveness of chemoprophylaxis with Maloprim administered by primary health care workers in preventing death from malaria amongst rural Gambian children aged less than five years old." Trans R Soc Trop Med Hyg 86(6): 580-1.

Piper, R., J. Lebras, et al. (1999). "Immunocapture diagnostic assays for malaria using Plasmodium lactate dehydrogenase (pLDH)." Am J Trop Med Hyg 60(1): 109-18.

Pittler, M. H. and E. Ernst (1999). "Artemether for severe malaria: a meta-analysis of randomized clinical trials." Clin Infect Dis 28(3): 597-601.

Price, R., M. van Vugt, et al. (1999). "Adverse effects in patients with acute falciparum malaria treated with artemisinin derivatives." Am J Trop Med Hyg 60(4): 547-55.

Rafael, M. E., T. Taylor, et al. (2006). "Reducing the burden of childhood malaria in Africa: the role of improved diagnostics." Nature 444 Suppl 1: 39-48.

Raharimalala, L. A., L. Rabarijaona, et al. (2002). "Malaria investigation in a cyclone-affected area in the south-east of Madagascar: entomological, diagnostic and therapeutic approaches." Archives de l'Institut Pasteur de Madagascar 68(1/2): 79-85.

Rasanen, P., E. Roine, et al. (2006). "Use of quality-adjusted life years for the estimation of effectiveness of health care: A systematic literature review." Int J Technol Assess Health Care 22(2): 235-41.

Reyburn, H., H. Mbakilwa, et al. (2007). "Rapid diagnostic tests compared with malaria microscopy for guiding outpatient treatment of febrile illness in Tanzania: randomised trial." Bmj 334(7590): 403.

Reyburn, H., R. Mbatia, et al. (2004). "Overdiagnosis of malaria in patients with severe febrile illness in Tanzania: a prospective study." Bmj 329(7476): 1212.

Reyburn, H., J. Ruanda, et al. (2006). "The contribution of microscopy to targeting antimalarial treatment in a low transmission area of Tanzania." Malar J 5: 4.

Rice, N. and A. Jones (1997). "Multilevel models and health economics." Health Econ 6(6): 561-75.

Rice, T. (1997). "Can markets give us the health system we want?" J Health Polit Policy Law 22(2): 383-426.

Rimón, M. M., S. Kheng, et al. (2003). "Malaria dipsticks beneficial for IMCI in Cambodia." Trop Med Int Health 8(6): 536-543.

Robert, V., K. Macintyre, et al. (2003). "Malaria transmission in urban sub-Saharan Africa." Am J Trop Med Hyg 68(2): 169-76

Robinson, R. (1993). "Cost-benefit analysis." Bmj 307(6909): 924-6.

Roca-Feltrer, A., I. Carneiro, et al. (2008). "Estimates of the burden of malaria morbidity in Africa in children under the age of 5 years." Trop Med Int Health 13(6).

Rolland, E., F. Checchi, et al. (2006). "Operational response to malaria epidemics: are rapid diagnostic tests cost-effective?" Trop Med Int Health 11(4): 398-408.

Russell, L. B. (2007). "Is all cost-effectiveness analysis local?" Med Decis Making 27(3): 231-2.

Sachs, J. (2002). Macroeconomics and health: Investing in health for economic development. . Geneva, World Health Organization.

Sachs, J. and P. Malaney (2002). "The economic and social burden of malaria." Nature 415(6872): 680-5.

Samuelson, W. and R. Zeckhauser (1988). "Status Quo Bias in Decision Making " J Risk and Uncertainty 1(1).

San Pedro, C. (1967/8). "Economic costs and benefits of malaria eradication." Phillipines J Pub Health 15: 5-24.

Schellenberg, J. A., J. Bryce, et al. (2004). "The effect of Integrated Management of Childhood Illness on observed quality of care of under-fives in rural Tanzania." Health Policy And Planning 19(1): 1-10.

Schultz, L. J., R. W. Steketee, et al. (1995). "Antimalarials during pregnancy: a cost-effectiveness analysis." Bull World Health Organ 73(2): 207-14.

Sculpher, M. J., F. S. Pang, et al. (2004). "Generalisability in economic evaluation studies in health care: a review and case studies." Health Technol Assess 8(49): iii-iv, 1-192.

Sendi, P., A. Gafni, et al. (2002). "Opportunity costs and uncertainty in the economic evaluation of health care interventions." Health Econ 11(1): 23-31.

Sendi, P. P. and A. H. Briggs (2001). "Affordability and cost-effectiveness: decision-making on the cost-effectiveness plane." Health Econ 10(7): 675-80.

Sheldon, T. A. (1996). "Problems of using modelling in the economic evaluation of health care." Health Econ 5(1): 1-11.

Shepard, D. S., M. B. Ettling, et al. (1991). "The economic cost of malaria in Africa." Trop Med Parasitol 42(3): 199-203.

Shiff, C. (2002). "Integrated approach to malaria control." Clin Microbiol Rev 15(2): 278-93.

Shillcutt, S., C. Morel, et al. (2008). "Cost-effectiveness of malaria diagnostic methods in sub-Saharan Africa in an era of combination therapy." Bull World Health Organ 86(2): 101-10.

Shillcutt, S., C. Morel, et al. (2007). Assessing RDT Cost-Effectiveness, WHO: http://www.wpro.who.int/sites/rdt/Assessing+RDT+Cost-Effectiveness.htm

Shillcutt, S. D., D. Walker, et al. (In press). "Cost-effectiveness in low- and middle-income countries: A review of the debates surrounding decision rules ". Pharmacoeconomics

Shrestha, R. K., B. Mugisha, et al. (2006). "Cost-effectiveness of including tuberculin skin testing in an IPT program for HIV-infected persons in Uganda." Int J Tuberc Lung Dis 10(6): 656-62.

Shretta, R., J. Omumbo, et al. (2000). "Using evidence to change antimalarial drug policy in Kenya." Trop Med Int Health 5(11): 755-64.

Shubhakaran and R. Jakhar (2004). "Cost effective treatment of acute, uncomplicated Plasmodium falciparum malaria." J Assoc Physicians India 52: 1009-10; author reply 1010.

Singh, N., A. Saxena, et al. (2002). "Usefulness of an inexpensive, Paracheck test in detecting asymptomatic infectious reservoir of Plasmodium falciparum during dry season in an inaccessible terrain in Central India." Journal of Infection 45(3): 165-168.

Sirima, S. B., A. Konate, et al. (2003). "Early treatment of childhood fevers with pre-packaged antimalarial drugs in the home reduces severe malaria morbidity in Burkina Faso." Trop Med Int Health 8(2): 133-9.

Snow, R. W. (2000). "The burden of malaria: understanding the balance between immunity, public health and control." J Med Microbiol 49(12): 1053-5.

Snow, R. W. and K. Marsh (1998). "New insights into the epidemiology of malaria relevant for disease control." Br Med Bull **54**(2): 293-309.

Sonnenberg, F. A. and J. R. Beck (1993). "Markov models in medical decision making: a practical guide." Med Decis Making **13**(4): 322-38.

Soto, J. (2002). "Health economic evaluations using decision analytic modeling. Principles and practices--utilization of a checklist to their development and appraisal." Int J Technol Assess Health Care **18**(1): 94-111.

Spiegelhalter, D. J., J. P. Myles, et al. (2000). "Bayesian methods in health technology assessment: a review." Health Technol Assess **4**(38): 1-130

Spiegelhalter, D. J., K. R. Abrams, et al. (2004). Bayesian Approaches to Clinical Trials and Health Care Evaluation. Chichester, JohnWiley &Sons.

Spiegelhalter, D. J., N. G. Best, et al. (2002). "Bayesian measures of model complexity and fit." J R Stat Soc **64**(4): 583-639

Spiegelhalter, D. J. and N. G. Best (2003). "Bayesian approaches to multiple sources of evidence and uncertainty in complex cost-effectiveness modelling." Stat Med **22**(23): 3687-709.

Ssengooba, F. (2004). "UGANDA'S MINIMUM HEALTH CARE PACKAGE: RATIONING WITHIN THE MINIMUM?" Health Policy and Development **2**(1).

Staedke, S. G., M. R. Kamya, et al. (2001). "Amodiaquine, sulfadoxine/pyrimethamine, and combination therapy for treatment of uncomplicated falciparum malaria in Kampala, Uganda: a randomised trial." Lancet **358**(9279): 368-74.

Stephens, J. K., K. Phanart, et al. (1999). "A comparison of three malaria diagnostic tests, under field conditions in North-west Thailand." Southeast Asian J Trop Med Public Health **30**(4): 625-30.

Stevens, W., V. Wiseman, et al. (2005). "The costs and effects of a nationwide insecticide-treated net programme: the case of Malawi." Malar J **4**(1): 22.

Stinnett, A. A. and J. Mullahy (1998). "Net health benefits: a new framework for the analysis of uncertainty in cost-effectiveness analysis." Med Decis Making **18**(2 Suppl): S68-80.

Sudre, P., J. G. Breman, et al. (1992). "Treatment of chloroquine-resistant malaria in African children: a cost-effectiveness analysis." Int J Epidemiol **21**(1): 146-54.

Susi, B., T. Whitman, et al. (2005). "Rapid diagnostic test for *Plasmodium* falciparum in 32 Marines medically evacuated from Liberia with a febrile illness." <u>Ann Intern Med</u> **142**(6): 476-7.

Swarthout, T. D., H. Counihan, et al. (2007). "Paracheck-Pf accuracy and recently treated *Plasmodium* falciparum infections: is there a risk of overdiagnosis?" <u>Malar J</u> **6**: 58.

Tan-Torres Edejer, T., R. Baltussen, et al. (2003). <u>Making Choices in Health: WHO Guide to Cost-Effectiveness Analyses</u>. Geneva, World Health Organization.

Tanser, F. C., B. Sharp, et al. (2003). "Potential effect of climate change on malaria transmission in Africa." <u>Lancet</u> **362**(9398): 1792-8.

Tediosi, F., N. Maire, et al. (2006). "An approach to model the costs and effects of case management of *Plasmodium* falciparum malaria in sub-saharan Africa." <u>Am J Trop Med Hyg</u> **75**(2 Suppl): 90-103.

Teerawattananon, Y. (2007). Assessing the Feasibility of Using Economic Evaluation in Reimbursement of Health Care Services in Thailand <u>School of Medicine, Health Policy and Practice</u> Norwich, University of East Anglia. **PhD.**

The East African Network for Monitoring Antimalarial Treatment (EANMAT) (2003). "The efficacy of antimalarial monotherapies, sulphadoxine-pyrimethamine and amodiaquine in East Africa: implications for sub-regional policy." <u>Trop Med Int Health</u> **8**(10): 860-867.

Tjitra, E., S. Suprianto, et al. (2001). "Detection of histidine rich protein 2 and panmalarial ICT Malaria Pf/Pv test antigens after chloroquine treatment of uncomplicated falciparum malaria does not reliably predict treatment outcome in eastern Indonesia." <u>Am J Trop Med Hyg</u> **65**(5): 593-8.

Trape, J. F. (2001). "The public health impact of chloroquine resistance in Africa." <u>Am J Trop Med Hyg</u> **64**(1-2 Suppl): 12-7.

Tversky, A. and D. Khanerman (1974). "Judgment under Uncertainty: Heuristics and Biases." <u>Science</u> **185**.

Uganda Ministry of Health (2005). IMPLEMENTATION GUIDELINES FOR THE HOME BASED MANAGEMENT OF FEVER TRATEGY IN CHILDREN .

UNICEF (2005). New Hope Against Malaria: Four Country Successes. UNICEF.

United Nations Population Division (2005). Population, Resources, Environment and Development: The 2005 Revision.

Utzinger, J., Y. Tozan, et al. (2001). "Efficacy and cost-effectiveness of environmental management for malaria control." Trop Med Int Health 6(9): 677-87.

Van Gool, K., G. Gallego, et al. (2007). "Economic Evidence at the Local Level : Options for Making it More Useful." Pharmacoeconomics 25(12): 1055-62.

Van Geertruyden, J. P., M. Mulenga, et al. (2006). "CD4 T-cell count and HIV-1 infection in adults with uncomplicated malaria." J Acquir Immune Defic Syndr 43(3): 363-7.

Van Geertruyden, J. P., M. Mulenga, et al. (2006). "HIV-1 immune suppression and antimalarial treatment outcome in Zambian adults with uncomplicated malaria." J Infect Dis 194(7): 917-25.

Von Seidlein, L., M. Jawara, et al. (2001). "Parasitaemia and gametocytaemia after treatment with chloroquine, pyrimethamine/sulfadoxine, and pyrimethamine/sulfadoxine combined with artesunate in young Gambians with uncomplicated malaria." Trop Med Int Health 6(2): 92-8.

Vounatsou, P. and T. Smith (1998). "Bayesian analysis of two component mixture distributions: application to the estimation of malaria attributable fractions." Appl Stat 47(4).

Waddington, C., J. Martin, et al. (2005). Trends in International Funding for Malaria Control. Prepared for the Roll Back Malaria Partnership. HLSP.

Wang, S. J., C. Christian Lengeler, et al. (2005). "Rapid urban malaria appraisal (RUMA) I: Epidemiology of urban malaria in Ouagadougou." Malar J 4(1): 43.

Warner, K. E. and R. C. Hutton (1980). "Cost-benefit and cost-effectiveness analysis in health care. Growth and composition of the literature." Med Care 18(11): 1069-84.

Wasunna, B., D. Zurovac, et al. (2008). "Why don't health workers prescribe ACT? A qualitative study of factors affecting the prescription of artemether-lumefantrine." Malar J 7: 29.

Weinstein, M. C., B. O'Brien, et al. (2003). "Principles of Good Practice for Decision Analytic Modeling in Health-Care Evaluation: Report of the ISPOR Task Force on Good Research Practices-Modeling Studies." Value in Health 6(1): 9-17.

White, N. J. (1983). "Severe hypoglycemia and hyperinsulinemia in falciparum malaria." N Engl J Med 309: 61-66.

White, N. J. (1992). "Antimalarial drug resistance: the pace quickens." J Antimicrob Chemother 30(5): 571-85.

White, N. J. (1999). "Delaying antimalarial drug resistance with combination chemotherapy."

White, N. J. (2005). "Intermittent presumptive treatment for malaria." PLoS Med 2(1): e3. Parassitologia 41(1-3): 301-8.

Whitty, C. J. and S. G. Staedke (2005). "Artemisinin-based combination treatment for malaria in Africa: no perfect solutions." Clin Infect Dis 41(8): 1087-8.

WHO (1996). "Investing in Health Research and Development: Report of the Ad Hoc Committee in Health Research Relating to Future Interventions Options". The World Health Organization. Geneva.

WHO (2001). "Antimalarial drug combination therapy: report of a WHO technical consultation". The World Health Organization. Geneva.

WHO (2002). "African Summit on Roll Back Malaria: The African Declaration and Plan of Action." The World Health Organization. Geneva.

WHO (2003). "Meeting Report: Malaria Rapid Diagnosis, Making it work." The World Health Organization. Geneva.

WHO (2003). The Abuja declaration and the plan of action. WHO/CDS/RBM. The World Health Organization. Geneva.

WHO (2004a). "Scaling up home-based management of malaria: from research to implementation". The World Health Organization. Geneva.

WHO (2004b). "The use of malaria rapid diagnostic tests." Roll Back Malaria, The World Health Organization. Geneva.

WHO (2006a). "Facts on ACTS - January 2006 update." The World Health Organization. Geneva.

WHO (2006b). "Guidelines for the treatment of malaria". The World Health Organization. Geneva.

WHO. (2006c). "WHO CHOICE - CHoosing Interventions that are Cost Effective."   Retrieved April 30th 2007, from http://www.who.int/choice/en.

WHO (2007a). Assessing RDT Cost-Effectiveness. The World Health Organization. Geneva.

WHO (2007b). Global AMDP database – SEARO. Retrieved June 13th 2008 from http://www.who.int/malaria/amdp/amdp_searo.htm.

WHO(2007c) World Health Statistics. The World Health Organization. Geneva

WHO (2008). CHOosing Interventions that are Cost-Effective: Country Specific Costs. Retrieved June 16th 2008 from http://www.who.int/choice/country/en/index.html#G

WHO-SIS (2007). World Health Statistics 2007, Tanzania Life Expectancy Tables, WHO. Retrieved May 12th 2008 from http://www3.who.int/whosis/life/life_tables/life_tables_process.cfm?path=whosis,life_tables&language=english

Wilkins, J. J., P. I. Folb, et al. (2002). "An economic comparison of chloroquine and sulfadoxine-pyrimethamine as first-line treatment for malaria in South Africa: development of a model for estimating recurrent direct costs." Trans R Soc Trop Med Hyg 96(1): 85-90.

Willan, A. R., D. Y. Lin, et al. (2005). "Regression methods for cost-effectiveness analysis with censored data." Stat Med 24(1): 131-45.

Williams, H. A., D. Durrheim, et al. (2004). "The process of changing national malaria treatment policy: lessons from country-level studies." Health Policy And Planning 19(6): 356-370.

Winstanley, P., S. Ward, et al. (2004). "Therapy of falciparum malaria in sub-saharan Africa: from molecule to policy." Clin Microbiol Rev 17(3): 612-37, table of contents.

Wiseman, V., W. A. Hawley, et al. (2003). "The cost-effectiveness of permethrin-treated bed nets in an area of intense malaria transmission in western Kenya." Am J Trop Med Hyg 68(4 Suppl): 161-7.

Wiseman, V., M. Kim, et al. (2006). "Cost-effectiveness study of three antimalarial drug combinations in Tanzania." PLoS Med 3(10): e373.

Wolfe, E. B., M. E. Parise, et al. (2001). "Cost-effectiveness of sulfadoxine-pyrimethamine for the prevention of malaria-associated low birth weight." Am J Trop Med Hyg 64(3-4): 178-86.

World Bank (1993). World Development Report. Washington DC.

World Bank. (2007). "GNI per capita 2005." Retrieved 23/05/07, from http://66.102.9.104/search?q=cache:00qbcXpOVb4J:siteresources.worldbank.org/DATAST ATISTICS/Resources/GNIPC.pdf.

World Bank (2007). GNI per capita 2005, Atlas method and PPP, World Bank.

Worrall, E., A. Rietveld, et al. (2004). "The burden of malaria epidemics and cost-effectiveness of interventions in epidemic situations in Africa." Am J Trop Med Hyg 71(2 Suppl): 136-40.

Ye, Y., C. Kyobutungi, et al. (2007). "Micro-epidemiology of *Plasmodium falciparum* malaria: Is there any difference in transmission risk between neighbouring villages?" Malar J 6: 46.

Yeka, A., G. Dorsey, et al. (2008). "Artemether-lumefantrine versus dihydroartemisinin-piperaquine for treating uncomplicated malaria: a randomized trial to guide policy in Uganda." PLoS ONE 3(6): e2390.

Yeung, S. (2006). Antimalarial Drug Resistance and Artemisinin Combination Threapy: A bio-economic model for the elucidation of policy Health Policy Unit. London, LSHTM. **Ph.D**.

Yeung, S., W. Pongtavornpinyo, et al. (2004). "Antimalarial drug resistance, artemisinin-based combination therapy, and the contribution of modeling to elucidating policy choices." Am J Trop Med Hyg 71(2 Suppl): 179-86.

Zurovac, D., B. A. Larson, et al. (2006). "The financial and clinical implications of adult malaria diagnosis using microscopy in Kenya." Trop Med Int Health 11(8): 1185-1194.

Zurovac, D., B. Midia, et al. (2006). "Microscopy and outpatient malaria case management among older children and adults in Kenya." Trop Med Int Health 11(4): 432-40.

Zurovac, D., J. Njogu, et al. (2008). "Effects of revised diagnostic recommendations on malaria treatment practices across age groups in Kenya." Trop Med Int Health 13(6): 784-7.

Zurovac, D., J. Njogu, et al. (2008). "Translation of artemether-lumefantrine treatment policy into paediatric clinical practice: an early experience from Kenya." Trop Med Int Health 13(1): 99-107.

Zurovac, D. and A. K. Rowe (2006). "Quality of treatment for febrile illness among children at outpatient facilities in sub-Saharan Africa." Ann Trop Med Parasitol 100(4): 283-96

Annex 1: Papers cited in literature review of economic analyses of malaria treatments/diagnostics

Trial based evaluations

Reference	Study summation	Cost perspective	Summary measure	Methodology
Honrado, Fungladda et al. 1999	RCT to determine cost-effectiveness of artesunate and QN + TTC. Follow up at 5 and 7 days as later parasitaemia could be due to reinfection or recrudescence and due to large proportion of losses to follow up.	Provider	Cost/expected number of patients cured	Simple CEA ratio
Gogtay, Kadam et al. 2003	Retrospective CEA comparing CQ, MFQ for the treatment of malaria in a referral centre in Mumbai. The data they used was gathered from a clinical trial in that particular hospital, and their initial stated aim is to estimate which of the treatments is most appropriate for use at that specific location	Provider	Average cost of treatment	Post-hoc simple CEA
Fernando, Karunaweera et al. 2004	A prospective study using immunochromatographic tests (ICTs) for the detection of Pv in a malaria endemic area of Sri Lanka. These were carried out alongside the routine use of blood slides, also used as the 'gold standard'. The authors report the findings for test sensitivity, specificity and predictive values, most significantly demonstrating a relatively low rate of sensitivity for ICTs for the detection of Pv.	Provider: Test costs	Cost/patient tested Cost/malaria case detected	Prospective study – all patients presenting tested with both RDT and microscopy.

Bualombai, Prajakwon g et al. 2003	Baulombai et al. report on a trial in Thailand comparing the use of two types of RDTs (detecting either pLDH or HRP2) with the use of microscopy. Their measure of effectiveness is test sensitivity, specificity, and its predictive values. Their analysis took a broader perspective to include patient costs.	Provider and consumer\n\nDiagnostic costs only	Cost/malaria case detected (by species)	Prospective study – all patients presenting tested with both RDT and microscopy.
Chanda et al. 2007	Evaluated the cost-effectiveness of AL vs SP in Zambia using data gathered from patients presenting at public health facilities in six district sites. Main outcome measures were treatment success and reduction in demand for second line treatment. Results suggest that AL produces successful treatment at less cost than SP, implying that AL is more cost-effective with an ICER of 4.10 US dollars per case successfully treated, and further cost savings when including 2nd line treatment costs.	Provider costs for drugs, labour, second line treatment	Average and incremental cost/case successfully treated	Prospective observational trial\n\nSimple CEA ratio
Wiseman et al. 2006	RCT carried out in a Tanzanian hospital to determine c/e of AQ ,AQ+SP,AQ+AS, AL. Almost uniquely in trial based evaluations captures both provider and patient costs	Provider and provider+patient	Total programme costs; average & incremental cost per case averted	Prospective study\n\nSimple CEA ratio
Jonkman Chibwe et al. 1995	Compared treatment strategies by conducting a trial based cost analysis for the implications of switching to a policy of microscopically confirmed parasitaemia prior to use of antimalarials in comparison to routine practice of PT	Community	Cost per treatment strategy	Cost analysis

Reference	Study summation	Cost perspective	Summary measures	Methodology
Pang and Piovesan-Alves, 2001	A cost-minimisation analysis on a community based program for the utilisation of RDTs and distribution of MFQ by bar staff in peripheral mining towns in Brazil.	Societal	Least cost	Cost minimisation analysis
Agnamey Brasseur et al. 2005	Evaluated the costs of switching from a policy of presumptive treatment with CQ or QN to treatment of microscopically confirmed patients with AS-AQ. This was done on a regional basis as a part of an evaluation of changing Senegalese national guidelines to such a policy.	Community	Day 28 sustained parasite clearance	CMA based on uncontrolled study of AS+AQ efficacy
Goodman, Coleman et al. 2000	Compared a range of malaria control interventions. These included preventive interventions - ITNs, residual spraying and prophylactic treatment, and interventions for improved case management such as improving compliance, changing first line drugs, the introduction of combination therapy and improving diagnostic practices.	Provider: Tests + malaria treatment costs	Cost/DALY averted For diagnosis net costs only Costs/test	Simple CEA, stratification by country income groups, transmission
Mulligan et al. 2005	Updated the previous evaluation, focusing on populations living in a high, stable transmission setting, in a low income SSA country. In terms of case-management the interventions included changing first line drugs for the treatment of uncomplicated malaria, with a focus on using of ACT.	Provider and community	Cost/ DALY averted	Modelling multiple secondary sources

Reference	Study summation	Cost perspective	Summary measures	Methodology
Cho Min, Lermaharit et al. 2000	DT model incorporating secondary data on treatment efficacies and costs for estimates of their cost-effectiveness, with 3 drug regimens as main choices (SP, CQ and Mefloquine), followed by chance nodes for efficacy (ACR, ETF, LTF), then by a chance node for compliance leading to terminal nodes (not/cured). Effectiveness was also measured by deaths prevented, taken as proportion of patients not cured. One way sensitivity analyses were conducted for treatment efficacies	Provider	Cost/malaria case cured Cost/malaria death prevented	Decision analysis model; one way sensitivity analyses
Sudre, Breman et al. 1992	The earliest attempt at conducting a decision analysis on the use of antimalarials, comparing CQ/SP/AQ for the treatment of children in SSA, under the circumstances of increasingly prevalent CQ resistant Pf. The model included a limited number of variables, illustrating the simplest of scenarios – a single treatment of febrile children, comparing the outcome for the different drugs in terms of cost per cure and cost per death averted.	Provider	Average and incremental costs for CQ,SP,AQ by levels of resistance to CQ	Decision analysis modelling
Wilkins, Folb et al. 2002	Detailed study of the average cost-effectiveness of SP and CQ. Using a range of data sources such as in vivo efficacy trials for the drugs in the location of interest (Mpumalanga, South Africa), diagnosis costs and patient travel costs, the authors constructed a decision tree to simulate comprehensive	Provider	Average cost-effectiveness ratio. Cost/case cured	Decision tree; Monte Carlo simulations

	treatment costs for each of the drugs. Treatment failure might lead to the return of the patient to the health care facility (probability assigned after consultation with local health workers) with either severe or uncomplicated malaria, followed by treatment with QN			
Muheki, McIntyre et al. 2004	Conducted an analysis of the cost-effectiveness of switching from SP to AL a resurgence of malaria in 1999 and 2000 required a number of interventions to bring it back under control, including reverting to the use of DDT in IRS programmes. Use of a Delphi survey amongst a number of experts estimated the impact of each of these interventions in reducing the incidence of malaria. With this estimate, along with cost data for the treatment, the authors were able to estimate the cost-effectiveness of the switch to AL.	Provider	Cost/strategy	Before/After study Provider
Rolland, Checchi et al. 2006	Compared the use of RDTs in malaria epidemics as a prerequisite to treatment with ACT. The authors used data collected in two such epidemics to populate a decision tree for a hypothetical population facing either presumptive treatment or a strategy of parasitological confirmation with RDTs.	Provider	Total costs; cost/false negative averted. COST-EFFECTIVENESS with WTP for FP averted	Decision tree modelling
Zurovac et al. 2006	A cost analysis Comparing 3 scenarios for different treatment guidelines and adherence to these for the treatment of patients with microscopy confirmed ACT at different levels of diagnostic accuracy.	Provider	Total cost per programme	Decision tree

Shillcutt, Morel et 2007	Decision tree analysis to estimate the cost-effectiveness of the use of RDT and microscopy in conjunction with ACT, in comparison to the predominant practice of presumptive treatment. In contrast to almost all other analyses, the authors also incorporate the health outcomes of non-malarial febrile illness (NMFIs)	Provider	Cost/DALY averted; probability of acceptance with varying WTP	Decision tree; Monte Carlo simulations; net benefit analysis
Rafael et al. 2007	Compared the use of diagnostic tests of varying accuracy across a range of levels of infrastructure (also a proxy measure for coverage). Only analysis to use an estimate of the harm associated with antimalarials. Used probabilistic sensitivity analysis to merge all parameter uncertainties (Not an economic evaluation – did not account for costs)	Societal (for health outcomes only)	Number of deaths. Number of unnecessary antimalarials used	Decision tree, probabilistic sensitivity analysis

Economic evaluations using System Dynamic models

Reference	Study summation	Cost perspective	Summary measures	Methodology
Schapira, Beals et al. 1993	Used a relatively simple mathematical model to estimate the costs and number of deaths averted with the use of a succession of different first line drugs in response to emerging resistance. The model suggests the durations for each of these that would result in lowest costs per death averted	Provider+cost of death	Minimum costs by timing of switch to new AM	Mathematical modelling
Laxminarayan, 2004	A mathematical bio-economic model to estimate malaria transmission, host immunity, and drug resistance to compare economic consequences of different treatment strategies (replacing CQ with ACT directly or first with SP and later ACT).	Societal	Strategy cost	Differential equation modelling

	The biological side of the model is comprised of calculations for rates of infection to resistant and susceptible strains, host immunity, mosquitoes per human and a number of other parameters, while the economics side accounts for productivity costs in addition to cost of treatment.			
Goodman, Coleman et al. 2001	Estimated the most appropriate point to switch from CQ to SP, modelling development of resistance and subsequent effectiveness of treatments if introduced at different points in time, and the costs this would incur. The model centres on a case-management DT for each regimen, with the outcomes at each chance node determining whether the patient is cured or not, followed by chance nodes for the patient's decision to pursue treatment with a second and third line drug in case of treatment failure. Costs and DALYs at each terminal node along with the probability of reaching each point could then provide an estimate of the cost-effectiveness of the regimen	Provider (+ cost of policy change)	Optimal year of switch; Cost per outpatient DALY averted	DT for CQ/SP with varying levels of resistance
Coleman, Morel et al, 2004	Estimate the incremental cost-effectiveness of introducing ACT, recognising that this must consider the temporal dynamics of drug resistance. As resistance changes and spreads, CE of treatment options will vary accordingly – of particular significance with ACT as resistance to these would be devastating to long term antimalarial efforts. Use of cost-effectiveness acceptability curves to inform decision makers of the probability that the use of ACT will be cost-effective, using different (subjective) thresholds	Provider + patient direct costs	Probability of switch to ACT being CE (<$150) at varying levels of initial R	Modelling levels of resistance; threshold analysis using DT
Morel, Lauer	Determined the cost-effectiveness of a range of interventions aimed at reducing the	Provider	Cost/Daly averted	Epidemiologica

et al. 2005	burden of malaria in SSA. These include treatment with CQ, SP and ACT, extending the use of ITNs, IRS and IPT. The design of the analysis is a generalised cost-effectiveness analysis, an approach advocated by WHO to facilitate generalisation of results and cross-regional comparisons.			I population model + Generalised CEA
Tediosi et al. 2006	Use a dynamic decision tree structure to capture the influence of treatment strategies on transmission intensity, looping back into the model to predict the incidence of clinical episodes and of mortality while incorporating effects of case management on persistence of parasites and transmission. Results reflect different levels of coverage across a 20 year time horizon	Societal	DALYS and YLLs lost Cost/capita/year	Stochastic simulation of Pf transmission to determine costs & health outcomes
Yeung, 2006	A bio-economic model that incorporated changes in resistance, transmission, and subsequent host-immunity to evaluate amongst other outcomes the cost-effectiveness of introducing ACT in place of monotherapy. The geographical focus of the research is a low transmission setting where resistance to existing antimalarials is high, although the model was also run to simulate high transmission intensities. Adherence and coverage were also central to the analysis, to ensure that it pertained to effectiveness and not limited to efficacy	Societal, provider	Cost/DALY averted	Bio-economic model; DT

Introduction

There is a surprising absence of data on the risk of progression from uncomplicated malaria to severe disease and death, particularly for untreated cases. Policy and clinical decisions are increasingly being guided by models, either of clinical algorithms or economic analyses. These models must include estimates of key parameters, but often data are lacking, and assumptions must be made based on little or no evidence, or by extrapolating data from different settings (e.g. South East Asia when commenting on Africa). Models of malaria often need to include information about non-malarial febrile illness since many depend on the outcome of antimalarial treatment, when in fact the problem is not malaria, but data on non-malarial febrile illnesses are also lacking.

In the absence of clear data, it seems sensible to investigate if there is a consensus opinion on the risk of severe illness and death, and if there is not a consensus, to see the range of opinions so that these estimates can be used in clinical and economic models. It seems a good moment to do this using a reasonably formal technique of which we can all take ownership collectively rather than any one group taking the lead. We suggest we do this by a Delphi survey. We have developed and piloted such a survey, which you will find below. For those of you who are interested in the technicalities of Delphi surveys or are not familiar with this method, there is additional information behind the questionnaire, but do not feel obliged to wade through it.

In brief, our aim is to circulate the survey in two rounds to experts in the field. In the first round, respondents will be asked to give their best estimates of the various key parameters which would be appropriate for use in economic and other models. We will then feed back the results of the first round to the same respondents and see if, in the light of the initial results, a consensus opinion can be achieved, or whether a wide-scatter of opinion continues.

The selection of panel members was purposive and entirely non-random, based on our own familiarity with potential panellists, their contribution to the literature and their clinical experience, aiming to recruit high profile individuals in the malariology community. We have chosen a relatively select group of people who have considerable experience in malaria and clinical management of patients based on their own immediate practice and knowledge of data from other centres. Participation in this Delphi survey will involve providing your opinion on how best to define the parameters of interest and your assessment of the parameter values.

When the Delphi survey is completed, we think it would be useful to submit this as a technical publication on which all respondents would be authors. The estimates derived from this survey will be useful for economic models of treatment and diagnostic practice, as well as other modelling, and will be freely available.

Round One Questionnaire

In this first round, we would like you to assign estimates for the probability that untreated malaria and febrile illness will progress to severe disease and death, for patients of different ages (< 5 years, 5-14 years, and > 15 years of age), and in areas of different malaria transmission intensity (hypoendemic, mesoendemic, hyper/holoendemic, guided by the definitions provided below). We would also like to gather opinions on whether these parameters are defined and stratified in the most appropriate manner, or how this might be modified. Please complete the following questionnaire by entering in your best estimate of the probabilities as a point estimate, and complete each question in full, if possible.

1. What is the probability that a patient with uncomplicated malaria (excluding pregnant women), who does not receive adequate treatment, will progress to severe malaria (any manifestation, including severe anaemia and cerebral malaria)?

In hypoendemic areas (EIR < 1; parasite prevalence < 10% in children aged 2-9 years):

      Age < 5 years: __ %

      Age 5-14 years: __ %

      Age > 15 years: __ %

In mesoendemic areas (EIR 1-100; parasite prevalence 11-50% in children aged 2-9 years):

    Age < 5 years: ___ %

    Age 5-14 years: ___ %

    Age > 15 years: ___ %

In hyper/holoendemic areas (EIR >100; parasite prevalence >50% in children aged 2-9 years):

    Age < 5 years: ___ %

    Age 5-14 years: ___ %

    Age > 15 years: ___ %

Additional comments:

<div style="border:1px solid black; height:80px;"></div>

2. What is the probability that a patient with severe malaria, who does not receive treatment, will progress to death?

In hypoendemic areas:

    Age < 5 years: ___ %

    Age 5-14 years: ___ %

    Age > 15 years: ___ %

In mesoendemic areas:

    Age < 5 years: ___ %

    Age 5-14 years: ___ %

    Age > 15 years: ___ %

In hyper/holoendemic areas:

     Age < 5 years: __%

     Age 5-14 years: __%

     Age > 15 years: __%

Additional comments:

| |
| |

3. Do you have additional suggestions on how the probabilities for untreated malaria listed in questions 1 and 2 above should be stratified, or what other risk factors should be considered? Please provide suggestions below:

| |
| |

4. What proportion of non-malarial febrile illnesses is likely due to bacterial illnesses that could be treated with antibiotics?

     Age < 5 years: __%

     Age 5-14 years: __%

     Age > 15 years: __%

5. What is the probability that non-malarial febrile illness, likely due to bacterial illness (including all possible infections, regardless of culture results), will become severe if not treated with antibiotics?

Age < 5 years: ___%

Age 5-14 years: ___%

Age > 15 years: ___%

6. What is the probability that severe non-malarial febrile illness, likely due to bacterial illness, will lead to death if not treated with antibiotics?

Age < 5 years: ___%

Age 5-14 years: ___%

Age > 15 years: ___%

7. Do you have additional suggestions on how non-malaria febrile illnesses should be classified, or the probabilities of progression to severe disease and death estimated? Please provide suggestions below:

8. Please indicate the geographic region from which you are basing your opinions:

9. Please provide any additional suggestions below:

Protocol and notes

Background to the survey

Economic evaluations based on decision models are routinely used to inform policy makers on choice of malaria diagnostics and treatments. Decision models are frameworks in which alternative courses of action are portrayed, and the superior option is identified based on the values of input parameters and the relationship between them. These models can be powerful tools, synthesizing a wide range of factors and producing clear decision recommendations. As such it is imperative that the best available evidence be used to inform parameter values.

Ideally all parameter values should be obtained from randomized control trials (RCTs), however, this is not always feasible or ethically viable. Such is the case with the health outcomes for patients with suspected malaria that do not receive correct diagnosis and appropriate treatment, which often serve as a baseline for evaluating the gains of malaria diagnostics and treatments. This data cannot be collected systematically in trials due to the obvious ethical considerations. Although expert opinion obtained via discussions with individual experts has been used to obtain estimates for these parameters [1-4], it is generally considered less methodologically sound than alternative sources of evidence, such as RCTs and observational studies [5, 6]. Regarding health outcomes for untreated malaria, expert opinions have varied considerably, and furthermore, they have not always made allowance for factors that can be highly influential, notably age and transmission intensity.

The differences in estimates for the probability of untreated malaria deteriorating to severe illness and death affect the comparability of evaluations in which they are used, and hence their usefulness to policy makers. Obtaining a consensus estimate for these parameters would encourage the use of an identical baseline in future evaluations and lend greater credibility to their results.

Delphi surveys are a well-established technique used to develop consensus on parameter values of interest. By eliciting the opinion of a range of highly-informed individuals, a consensus on parameter values of interest is distilled. The process invites input from a large number of

246

individuals through a systematic, anonymous and iterative feedback dynamic between group members. The ultimate aim of consensus building is to minimize the variance around parameter values [7]. The approach facilitates a more inclusive and democratic process as opposed to open discussions where a small number of individuals can dominate discussion and consequent opinion [8].

Despite the existence of a large number of uncertain parameters, use of Delphi surveys in the context of malaria treatment and diagnostics has been limited. Sudre, Breman and Koplan used a Delphi survey when evidence on CQ resistant was beginning to emerge in order to obtain probabilities of treatment failure amongst children of different age groups [9]. Muheki, McIntyre and Barnes used a Delphi survey to assess the contribution of ACT usage to the reduction in malaria transmission in KwaZulu Natal [10].

This study endeavours to obtain estimates for a number of baseline parameters, reflecting the probabilities of patients with untreated malaria and non-malarial febrile illnesses developing severe illness, and the consequent case-fatality rates for these. These parameters would be used to help evaluate the comparative effectiveness and efficiency of alternative malaria control interventions.

Survey aims and objectives

This study aims to construct a set of probabilities relating to the transition from untreated, uncomplicated malaria and malaria-like febrile illnesses to severe illness and death. The first round questionnaire aims at identifying the relevant parameters for estimation in subsequent rounds and to obtain initial estimates for these. In subsequent rounds participants revise these estimates to distil a clear range of opinions and where possible attain a consensus on these.

There are two probability sets to be modified and estimated, relating to malaria and non-malarial febrile illness (NMFI). Both malaria and NMFIs are included in the study because malaria control interventions often have an impact on the management of febrile episodes as a whole, many of which have other causes of illness [11-13]. The outcomes of NMFIs are increasingly accounted for in the evaluation of malaria treatment and diagnostics and estimates for their prognosis are equally as important [3, 4, 14].

Regarding the probability for untreated malaria becoming severe and causing death, a number of stratifications will likely be considered necessary. The questionnaire has initially been designed to stratify by transmission intensity and age [15,16]. Other factors such as HIV prevalence might also be suggested as necessary. The first round questionnaires will aim to elicit participants' opinions on the factors considered most relevant and methods used to differentiate between their different strata. The main anticipated challenge to be conveyed to panellists is ensuring the number of stratifications is kept at a minimum to avoid the probability set from becoming too unwieldy[13].

The second required probability set will deal with NMFIs. There are a number of challenges to estimating the probability for these progressing to severe illness. Firstly there is very little data on the breakdown of these illnesses and there is likely to be considerable variation in their composition, dependent on factors such as age and location. Secondly, in the context of decision modelling, trying to capture all possible illnesses in the models will be impractical. For these reasons a broader classification is required. Panellists will therefore be consulted on whether NMFIs can be classified based on whether or not the illnesses require treatment with antibiotics. This classification assumes that malaria like illnesses that do not require antibiotics (or antimalarials) are by and large self-limiting. Transition probabilities are then sought only for illnesses that would require antibiotics.

Overview of the method

Delphi surveys use a series of questionnaires distributed amongst participants in a number of rounds with the aim of achieving a consensus on an answer to a research problem. The choice of participants is a selective one, based on the participant's familiarity with the topic and

---

[13] Each additional factor such as HIV prevalence multiplies the number of estimates required by the number of strata, so that for instance 3 transmission intensities, with three age groups and three levels of HIV prevalence will require 27 assessments for the probability of a particular age-TI-HIV configuration developing severe malaria when untreated.

representative sampling is not appropriate [17]. There is strong support for ensuring a diverse range of opinions spanning the full spectrum of respectable controversy both to ensure accuracy and lend credibility to results [18].

The first round questionnaire is aimed primarily at constructing appropriate questions and definitions for panellists to address in later rounds. It allows for considerable qualitative as well as quantitative responses which are summarised and fed back to participants. In subsequent rounds, participants will be able to enter their responses and then revise their estimates iteratively after reflecting on previous results and other participants' arguments. In the earlier days of Delphi surveys four rounds were considered ideal, although in more recent studies two or three rounds have been accepted as sufficient. The decision on the number of rounds tends to be a pragmatic one [17, 18].

Presentation of interim results at each round is controlled by the facilitators, analysing and presenting qualitative responses from the first round and summarizing quantitative results for subsequent rounds. Use of measures of dispersion as well as central tendency is encouraged in order to demonstrate the range of opinions and phenomena such as clustering around divergent estimates. There is no firm rule to determine when a consensus is reached; this is indicated primarily through a reduction in variance [18].

It should be noted that there are a number of potential biases in Delphi surveys. The selection of panellists tends to determine the range and nature of views expressed in the surveys [19]. The processing and feedback of interim results, particularly of qualitative responses, is managed by the facilitator and subject to their own prioritization. Panellists who feel their estimates diverge significantly from those reflected in subsequent rounds could withdraw from the process.

While these and other potential biases have drawn considerable criticism, Delphi surveys retain credibility where they are as transparent as possible and demonstrate a clear *decision trail* describing and justifying choices made at all stages [18].

Use of Results

The survey findings are intended to be used primarily in decision models concerned with evaluating the costs and benefits of malaria diagnostics and treatment. The survey is expected to produce both measures of central tendency to be used as point estimates, and also distributions that can be used as a basis for sensitivity analyses, reflecting how the uncertainty surrounding probabilities influences results.

Survey results will be conveyed to all panellists and made freely available for use in evaluations of malaria control interventions through publication in a peer reviewed journal, which all panellists will be invited to co-author.

Further information

Any queries on the survey or on filling in the questionnaire can be sent to Yoel.Lubell@lshtm.ac.uk. A hard copy of this document can be sent upon request.

References

[1]   Akhavan D, Musgrove P, Abrantes A, d AGR. Cost-effective malaria control in Brazil. Cost-effectiveness of a Malaria Control Program in the Amazon Basin of Brazil, 1988-1996. Social science & medicine (1982). 1999 Nov; 49(10):1385-99.

[2]   Goodman CA, Coleman PG, Mills AJ. Changing the first line drug for malaria treatment--cost-effectiveness analysis with highly uncertain inter-temporal trade-offs. Health Econ. 2001 Dec; 10(8):731-49.

[3]   Lubell Y, Reyburn H, Mbakilwa H, Mwangi R, Chonya S, Whitty CJ, et al. The cost-effectiveness of parasitological diagnosis for malaria-suspected patients in an era of combination therapy. Am J Trop Med Hyg. 2007; 77(Suppl 6):128-32.

[4]   Shillcutt S, Morel C, Goodman C, Coleman P, Bell D, Whitty CJM, et al. Cost-effectiveness of malaria diagnostic methods in sub-Saharan Africa in an era of combination therapy. Bulletin of the World Health Organization. 2008; 86(2):101-10.

[5]   Evans D. Hierarchy of evidence: a framework for ranking evidence evaluating health care interventions. Journal of clinical nursing. 2003 Jan; 12(1):77-84.

[6]   Haynes RB. What kind of evidence is it that Evidence-Based Medicine advocates want health care providers and consumers to pay attention to? BMC Health Serv Res. 2002; 2(1):3.

[7]   Graham B, Regehr G, Wright JG. Delphi as a method to establish consensus for diagnostic criteria. Journal of clinical epidemiology. 2003 Dec;56(12):1150-6.

[8]   Butterworth T, Bishop V. Identifying the characteristics of optimum practice: findings from a survey of practice experts in nursing, midwifery and health visiting. J Adv Nurs. 1995 Jul;22(1):24-32.

[9]   Sudre P, Breman JG, Koplan JP. Delphi survey of malaria mortality and drug resistance in Africa. Lancet. 1990 Mar 24; 335(8691):722.

[10]   Muheki C, McIntyre D, Barnes KI. Artemisinin-based combination therapy reduces expenditure on malaria treatment in KwaZulu Natal, South Africa. Trop Med Int Health. 2004 Sep; 9(9):959-66.

[11]   Amexo M, Tolhurst R, Barnish G, Bates I. Malaria misdiagnosis: effects on the poor and vulnerable. Lancet. 2004 Nov 20; 364:1896-8.

[12]   Reyburn H, Mbakilwa H, Mwangi R, Mwerinde O, Olomi R, Drakeley C, et al. Rapid diagnostic tests compared with malaria microscopy for guiding outpatient treatment of febrile illness in Tanzania: randomised trial. Bmj. 2007 Feb 24; 334(7590):403.

[13]   Tarimo D, Minjas J, Bygbjerg I. Malaria diagnosis and treatment under the strategy of the integrated management of childhood illness (IMCI): relevance of laboratory support from the rapid immunochromatographic tests of ICT Malaria P.f./P.v. and OptiMal. Ann Trop Med Parasitol. 2001; 95(5):437-44.

[14]   Rafael ME, Taylor T, Magill A, Lim YW, Girosi F, Allan R. Reducing the burden of childhood malaria in Africa: the role of improved diagnostics. Nature. 2006 Nov 23; 444 Suppl 1:39-48.

[15]   WHO (2008). Global Malaria Control and Elimination: report of a technical review. T. W. H. Organization. Geneva. **January 17-18, 2008**.

[16]   Marsh, K. and R. W. Snow (1999). "Malaria transmission and morbidity." Parassitologia **41**(1-3): 241-6.

[17]   Hasson F, Keeney S, McKenna H. Research guidelines for the Delphi survey technique. J Adv Nurs. 2000 Oct; 32(4):1008-15.

[18]   Powell C. The Delphi technique: myths and realities. J Adv Nurs. 2003 Feb; 41(4):376-82.

[19]   Keeney S, Hasson F, McKenna H. Consulting the oracle: ten lessons from using the Delphi technique in nursing research. J Adv Nurs. 2006 Jan; 53(2):205-12.

Annex 3: Cost-effectiveness of artesunate for the treatment of severe malaria[14]

Y Lubell, S Yeung, AM Dondorp, NP Day, F Nosten, E Tijtra, Md Abul Faiz, E Bin Yunus, NM Anstey, SK Mishra, S Mohanty, NJ White and AJ Mills

Objective: Artesunate has been shown to be superior to quinine for the management of severe malaria, in clinical trials conducted in Asia. This study explores the cost-effectiveness of artesunate based principally on the findings of a large multi-centre trial carried out in Southeast Asia.

Methods: Trial data were used to compare mortality amongst patients with severe malaria, treated with either artesunate or quinine. This was combined with retrospectively collected cost data to estimate the incremental cost per death averted with the use of artesunate instead of quinine.

Results: The incremental cost per death averted using artesunate was approximately 140USD. Artesunate maintained this high level of cost-effectiveness also when allowing for the uncertainty surrounding the cost and effectiveness assessments.

Conclusion: This analysis confirms the vast superiority of artesunate for treatment of severe malaria from an economic as well as a clinical perspective.

Word count: Abstract – 141. Main text – 2513.

---

[14] In press, *Tropical Medicine and International Health*

Background

Artemisinin combination therapies (ACT) are now recommended as first line treatment of uncomplicated falciparum malaria in almost all malaria endemic countries [1]. They have been shown repeatedly to be more effective and cost-effective than their predecessors [2-5], require only once a day dosing and are associated with few adverse effects. For severe malaria, quinine has been the traditional standard treatment in both developed and developing countries. Quinine is an effective antimalarial, but it is not simple to administer, and it has a narrow therapeutic ratio. It is associated with a risk of local toxicity following intramuscular injection, and significant risks of systemic toxicity (hypoglycaemia, hypotension if administered rapidly). Quinine must be given three times daily either by rate-controlled intravenous infusion or intramuscular injection to the anterior thigh [6]. A growing body of evidence summarised in recent reviews demonstrates the considerable superiority of artesunate relative to quinine in terms of mortality rates without increasing rates of neurological deficit (Cochrane review estimate; RR 0.62, 95% CI 0.51 to 0.75 [7]). The studies so far have included mostly adults in Asia, although of the 1461 patients enrolled into the large multi-centre SEAQUAMAT trial, 202 were children, and benefits were similar in both age-groups [8].

The SEAQUAMAT study was conducted across ten sites in four South East Asian countries. A large multi-centre trial of similar design is underway to estimate the comparative effectiveness of artesunate in children in Sub-Saharan Africa. In the SEAQUAMAT trial, mortality in patients treated with artesunate was 35% lower than in quinine recipients. The implication was that, on average, for every 13 patients treated with artesunate instead of quinine, one death would be averted.

Despite these promising results, and endorsement by the WHO treatment guidelines, many local guidelines in malaria endemic countries continue to recommend quinine as the drug of choice for severe malaria [9]. The second most frequently recommended treatment for severe malaria is artemether [1], even though its advantage over quinine in terms of mortality has been shown to be limited [10, 11]. Artesunate has only recently been added to the policy guidelines of a limited number of countries in Asia [1], and its cost-effectiveness has yet to be assessed.

The aim of this paper is therefore to examine the costs and consequences of switching from quinine to artesunate from an economic perspective.

Methodology

A cost-effectiveness analysis framework was used to determine the cost per death averted by switching from quinine to artesunate for inpatients with severe malaria.

Interventions. The interventions being considered were quinine and artesunate for treatment of severe malaria. The drugs were given intravenously. Once patients had recovered sufficiently to take tablets, they continued with the same antimalarial taken orally to complete a course of 7 days.

Perspective. The perspective taken was that of the provider, so only costs incurred by the hospitals were accounted for, as this was considered of most immediate relevance for decision making purposes by ministries of health considering policy change.

Trial data. The SEAQUAMAT study was a multi-centre trial carried out between 2003-2005 in one site each in Bangladesh, India, Indonesia and seven sites in Myanmar and has been described in detail elsewhere [8]. Relevant patient-specific data from the trial for this analysis include mortality for each of the drugs, dosages used, and the length of stay in hospital as inpatients. The incidence of significant neurological sequelae are also summarised although as their incidence in the prospective studies was very low, they were not incorporated in the final measure of outcome.

Cost data. The costs included are the provider costs resulting from a switch from quinine to artesunate. Costs for artesunate were obtained from the producer and include shipment costs. Quinine costs and those for i.v. sets and syringes to administer the drugs were obtained from the International Drug Price Indicator Guide [12]. The cost of i.v. sets and syringes used to administer the drugs are presented for general comparison, however they are not included in the calculation of the cost per death averted due to the variability in routine administration practices of the treatments. Drug costs were increased by 15% to account for taxes and an

254

extra 10% for wastage [13]. Standard inpatient care costs for each country were obtained from the WHO-CHOICE database and are specific to the level of hospital at each site. These costs include "hotel" expenditures – those for personnel, capital and food and exclude drug costs [14]. Both drug costs and those for inpatient care were calculated for each patient individually based on their length of stay and drug dosage used.

It was assumed that, apart from the cost of trial drugs, the inpatient costs per day were the same for both treatment arms. Labour costs were also assumed to be equal, despite the fact that artesunate is considerably easier and simpler to administer. Costs were converted from local units to US dollars at the relevant year, adjusted for inflation using the consumer price index, and reported in 2008USD. Table 1 shows the unit costs used.

Item	Unit cost (quantity)		Source and notes
Quinine vial (300mg salt)	$0.19 (1)		International Drug Price Indicator Guide. Accessed 2/6/08
Quinine tab (300mg)	$0.04 (1)		
Artesunate vial (60mg)	$1.2 (1)		Quote from the producer, Guilin Pharma, Personal communication , 16/5/08
Artesunate tab (50mg)	$0.17 (1)		
Artesunate administration equipment	$0.3 (1x 5ml syringe and 2 x needles)		
Quinine administration equipment	$1.2 (1x 5 ml syringe, 2 x needles, 1 x infusion set, 1 IV solution)		
Cost per inpatient day	Bangladesh	$5.7	WHO-CHOICE estimates by country and hospital level [14]. Accessed 23/5/08
	India	$10.25	
	Indonesia	$2.5	
	Myanmar	$2.1 (mean)	

Table 5: Costs for treatment, equipment and inpatient care used in the analysis

Analysis. The number needed to treat (NNT) to avert one death was calculated from the difference in mortality between the two arms. This was multiplied by the difference in average cost of treating a patient with each of the drugs, providing the incremental cost per death averted using artesunate. This figure is equivalent to the incremental cost-effectiveness ratio

255

(ICER) calculated using standard methods [15]. The primary outcome is the pooled estimate for cost per death averted, merging cost and effectiveness data from all sites. As the trial was a multi-centre study, results are also reported stratified by country.

Discounting. The immediacy of the costs and benefits (deaths averted) meant that no discounting was needed.

Sensitivity analysis. Uncertainty surrounding both outcomes and costs was explored using probabilistic sensitivity analysis. The treatment outcome, a binary variable with values representing outcomes of either *dead* or *alive*, was assigned a beta distribution using the mortality frequency to define the distribution parameters. This allows for the greater uncertainty in sites that recruited fewer patients. Probability distributions were fitted to the cost of antimalarials given to all patients (gamma distributions; these are skewed to the right reflecting the tendency of cost data to be positively skewed [16]). Using the @Risk Excel plug-in, a Monte Carlo simulation was carried out to observe the impact these uncertainties had on the ICER.

A threshold analysis was carried out to estimate the cost at which artesunate ceases to be cost-effective. This was done using decision thresholds of $575 and £3450 per death averted. These values were obtained using WHO's thresholds for the willingness to pay to avert the loss of a disability adjusted life year [17,18], multiplied by the average remaining life expectancy for a patient that survives their illness, based on life expectancy tables for the relevant countries [19] , and discounted at 3% [13].

Results

Health outcomes are shown in Table 2. The mortality amongst patients treated with artesunate was considerably lower in all countries; the pooled estimate for mortality amongst the artesunate group was 34.7% lower than that for quinine (95% CI 18.5-47.6%; p=0.0002). There were very few instances of neurological sequelae, 3 in the quinine arm and 7 in the artesunate arm. The difference between these was not statistically significant (p=0.2).

Country		Bangaladesh	India	Indonesia	Myanmar	Pooled
Quinine	Mortality	**32.5%** (75/231)	**26.4%** (19/72)	**12.3%** (18/146)	**18.4%** (52/282)	**22.4%** (164/731)
	Neurological sequelae	0	0	1	2	3
Artesunate	Mortality	**23.4%** (52/222)	**21.4%** (15/70)	**6.3%** (9/143)	**10.5%** (31/295)	**14.7%** (107/730)
	Neurological sequelae	3	1	1	2	7

Table 6: Patient outcomes by country and pooled

The mean costs for each country and the pooled estimate are shown in Table 3. The variation in costs is due to both differences in unit costs, and in the average length of stay in each hospital. Despite this variation, overall the difference in average cost between the treatment groups was fairly consistent, costs being slightly higher for recipients of artesunate.

	Bangaladesh	India	Indonesia	Myanmar	Pooled
Quinine					
Drug cost	$5.0	$5.5	$4.6	$4.8	$4.8
Drug administration cost	$8.2	$9.3	$9.0	$8.5	$8.7
Inpatient care cost	$21.8	$45.6	$50.2	$12.7	$27.6
Total	**$35.0**	**$60.4**	**$63.8**	**$26.0**	**$41.1**
Artesunate					
Drug cost	$15.8	$21.6	$13.4	$13.0	$14.7
Drug administration cost	$1.1	$1.4	$1.0	$1.0	$1.1
Inpatient care cost	$27.0	$47.3	$52.6	$12.9	$28.3
Total	**$44.0**	**$70.3**	**$67.0**	**$26.9**	**$44.1**

Table 7 - Treatment cost per patient by country. IPcost - cost per inpatient day excluding drugs

Costs and outcomes are combined in table 4. Given the number of assumptions involved, the costs in this table exclude the equipment needed to administer the drugs, making artesunate appear even more costly than quinine compared to the totals in Table 3. As previously reported [8], treatment with artesunate was associated with a relative risk of 0.65. The bottom row specifies the incremental cost per death averted. The pooled ICER is $135.6, with a range of values from $104 in Myanmar where the largest number of patients were recruited, to $361 in India where fewest patients were recruited.

Country	Bangaladesh	India	Indonesia	Myanmar	Pooled
Mean cost for patients treated with quinine	$26.8	$51.1	$54.8	$17.5	$32.4
Mean cost for patients treated with artesunate	$42.8	$68.9	$66.0	$25.8	$43.0
Relative risk for treatment with artesunate	0.72	0.81	0.51	0.57	0.65
Numbers needed to treat to avert a death	11	20	17	13	13
Incremental cost per death averted	**$177.2**	**$358.9**	**$185.7**	**$104.8**	**$135.6**

Table 8 – Costs for each of the treatments, combined with the relative risk to produce the numbers needed to treat and the incremental cost per death averted.

Sensitivity analysis

By assigning a probability distribution to all hospital inpatient care costs and treatment outcomes, the uncertainty surrounding these parameters is carried through to the ICER for artesunate. Using a Monte Carlo simulation, the mean cost per death averted was found to be $140.2, approximating the deterministic calculation. By removing the highest and lowest 2.5% of observations, a 95% interval was created, ranging from -$120 to $455. The negative values indicate those instances where artesunate is more effective and less costly, so providing hospitals with savings in costs of treating severe malaria[15].

Use of the threshold analysis to estimate the point at which artesunate ceases to be cost-effective showed that if decision makers are willing to pay $575 to avert the loss of a life, artesunate would be cost-effective up to a cost of $4.2 per vial, over three times its current selling price. For a higher threshold of $3450 the cost could be $24.4, over 20 times its current selling price, before the drug ceases to be cost-effective.

---

[15] Although a negative value could also imply the inverse (artesunate associated with higher mortality and lower costs than quinine), in this case artesunate was found to be superior in all iterations of the simulation.

Discussion

This economic evaluation confirms the considerable superiority of artesunate over quinine for the treatment of severe malaria, at least for adults in South East Asia. The pooled estimate suggests that the incremental cost of averting a death using artesunate is approximately $140. The variation in results both by country and in the sensitivity analysis do not diminish the strength of these results. In fact the sensitivity analysis shows that the use of artesunate can be cost saving, in addition to being clinically superior.

Cost-effectiveness analyses require a comparison of results to decision thresholds to determine whether an intervention can be considered a good investment. GDP per capita is increasingly considered a benchmark for determining when an intervention is considered cost effective, with GDP per capita being compared to the cost of averting the loss of a life year in full health [20]. As the mean cost per death averted is well below the GDP per capita of even the poorest countries, there is no doubt that the use of artesunate to treat severe malaria represents an extremely good investment. This is further supported by the fact that these cost estimates suggest a much better return on investment than those for other well accepted malaria related interventions, such as $858 per death averted from implementing an environmental control programme [21], or a range of $254 to $3437 per death averted for insecticide treated mosquito nets [22].

A number of simplifying assumptions have been made that reduce the advantage of artesunate. First no attempt has been made to quantify and cost the labour requirements of administering the drugs. Quinine is administered three times per day and infusions need careful monitoring whereas artesunate is given by as a daily bolus injection and therefore requires no special nursing attention. Second, the costs for equipment used to administer the drugs, shown to be higher for quinine, were not included in the calculation of cost per death averted. This was done as quinine can be given by the intramuscular as well as the intravenous route, and patients with severe illness will often be given intravenous fluids irrespective of the route of drug administration. Although not used in the calculation of cost per death averted, the estimate of the different costs of administering the drugs is provided in table 3 and clearly favours artesunate. Third, any potential differences in the treatment of adverse events such as

259

hypoglycaemia, a specific adverse effect of quinine [23], have also been ignored but would obviously favour artesunate. In the trial hypoglycaemia following treatment was indeed less frequent with artesunate (Mantel-Haentzel stratified odds ratio and 95%CI from the SEAQUAMAT study; 0·31 [0·12–0·78]), although monitoring for hypoglycaemia would still be required for both treatments [6].

Despite these assumptions, that all reduce the potential cost-effectiveness of artesunate, the cost per death averted demonstrates that it remains a highly attractive intervention. The production of a GMP artesunate is currently underway[16] and as the sensitivity analysis shows, artesunate would continue to be cost-effective for the treatment of severe malaria even if its price was significantly higher.

Neurological sequelae were reported but not accounted for as these were found not to differ significantly and occurred in less than 1% of cases, although importantly there was no evidence of a significant increase in risk in the artesunate group (7/730) compared with the quinine group (3/731). Reporting the results as cost per Disability Adjusted Life Year (DALY) averted would therefore add little to the analysis.

The perspective taken in this analysis is that of the provider. Ideally, economic evaluations would include a broader range of costs and benefits, including costs incurred by patients, and how these differ between interventions. In this instance there was no reason to expect major differences in costs, and as the data for this were not readily available, this was excluded from the analysis.

Multi-centre trials can pose a number of challenges in analysing and interpreting results. Pooling data is not always a valid procedure while stratifying results by site may result in significant loss of power and fail to make full use of available data [24]. Pooling cost data raises concerns around how prices are standardized, as unit costs can vary widely across countries. In this

---

[16] Safety and Efficacy Study of IV Artesunate to Treat Malaria

http://clinicaltrials.gov/ct2/show/NCT00298610 Accessed November 11th 2008

analysis prices were standardised using official exchange rates and inflating their value from the year in which they were reported to 2008 USD. An alternative approach is the use of international dollars that adjust for purchasing power parity, however this has not yet entered mainstream use and can cause confusion for those not familiar with purchasing power adjustment.

Some variation in effectiveness between sites was observed, but as there was no prior reason to expect significant differences in treatment effects, and the observed differences were not statistically significant, pooling the data was considered justifiable.

Conclusion

With over a million annual deaths due to malaria, it is imperative that the most effective and cost-effective treatments be used for patients with severe illness. Artesunate is considerably superior to quinine for the treatment of severe malaria in Asia and has been incorporated into the national guidelines in a limited number of countries including China, Vietnam, India, Indonesia, Thailand, Lao PDR and PNG. However the guidelines in many countries in the region continue to recommend quinine or artemether. This study has demonstrated that from a cost-effectiveness perspective, substituting quinine with artesunate would provide a return on investment that few health interventions could match in terms of both immediate health gains and minimal, if any, additional cost. There seems no reason to deny patients the best available treatment.

References

1.   WHO, *Global AMDP database - SEARO*. 2007, http://www.who.int/malaria/amdp/amdp_searo.htm. Accessed 27/6/08
2.   Agnamey, P., et al., *Economic evaluation of a policy change from single-agent treatment for suspected malaria to artesunate-amodiaquine for microscopically confirmed uncomplicated falciparum malaria in the Oussouye District of south-western Senegal.* Trop Med Int Health, 2005. **10**(9): p. 926-33.
3.   Honrado, E.R., et al., *Cost-effectiveness analysis of artesunate and quinine + tetracycline for the treatment of uncomplicated falciparum malaria in Chanthaburi, Thailand.* Bull World Health Organ, 1999. **77**(3): p. 235-43.

4. Wiseman, V., et al., *Cost-effectiveness study of three antimalarial drug combinations in Tanzania*. PLoS Med, 2006. **3**(10): p. e373.

5. Yeung, S., *Antimalarial Drug Resistance and Artemisinin Combination Threapy: A bio-economic model for the elucidation of policy* in *Health Policy Unit*. 2006, LSHTM: London.

6. Anstey, N.M., R.N. Price, and N.J. White, *Improving the availability of artesunate for treatment of severe malaria*. Med J Aust, 2006. **184**(1): p. 3-4.

7. Jones, K.L., S. Donegan, and D.G. Lalloo, *Artesunate versus quinine for treating severe malaria*. Cochrane Database Syst Rev, 2007(4): p. CD005967.

8. Dondorp, A., et al., *Artesunate versus quinine for treatment of severe falciparum malaria: a randomised trial*. Lancet, 2005. **366**(9487): p. 717-25.

9. WHO, *Guidelines for the treatment of malaria*. 2000. The World Health Organization: Geneva.

10. Hien, T.T., et al., *A Controlled Trial of Artemether or Quinine in Vietnamese Adults with Severe Falciparum Malaria*. New England Journal of Medicine, 1996. **335**(2): p. 76.

11. Pittler, M.H. and E. Ernst, *Artemether for severe malaria: a meta-analysis of randomized clinical trials*. Clin Infect Dis, 1999. **28**(3): p. 597-601.

12. MSH., *International Drug Price Indicator Guide*. 2007, http://erc.msh.org/dmpguide.

13. Gold, M.R., et al., *Cost Effectiveness in Health and Medicine*. 1996, New York: Oxford University Press.

14. WHO, *CHOosing Interventions that are Cost-Effective: Country Specific Costs*. 2008.

15. Drummond, M.F., et al., *Methods for the Economic Evaluation of Health Care Programmes*. Third ed. 2005: Oxford University Press.

16. Nixon, R.M. and S.G. Thompson, *Parametric modelling of cost data in medical studies*. Stat Med, 2004. **23**(8): p. 1311-31.

17. Jamison, D. T., W. H. Mosely, et al. (1993). Disease control priorities in developing countries. Oxford, Oxford University Press.

18. WHO CHOICE - CHoosing Interventions that are Cost Effective." Retrieved April 30th 2007, from http://www.who.int/choice/en

19. WHOSIS – Life Tables for WHO Member States. Retrieved November 12th 2008 from http://www.who.int/whosis/database/life_tables/life_tables_process.cfm?path=whosis,l ife_tables&language=english

20. Sachs, J., *Macroeconomics and health: Investing in health for economic development.* . 2002, World Health Organization: Geneva.

21. Utzinger, J., Y. Tozan, and B.H. Singer, *Efficacy and cost-effectiveness of environmental management for malaria control*. Trop Med Int Health, 2001. **6**(9): p. 677-87.

22. Mulligan, J., C. Morel, and A. Mills, *The Cost-Effectiveness of Malaria Control Interventions*, in *DISEASE CONTROL PRIORITIES PROJECT*. 2005.

23. White, N.J., *Severe hypoglycemia and hyperinsulinemia in falciparum malaria*. N Engl J Med, 1983. **309**: p. 61-66.

24. Grieve, R., et al., *Multilevel models for estimating incremental net benefits in multinational studies*. Health Econ, 2007. **16**(8): p. 815-26.