Rapid diagnostic tests for malaria: effect on quality of care under experimental conditions and in routine practice

Thesis submitted in accordance with the requirements of the University of Liverpool for the degree of Doctor in Philosophy

John Odaga

April 2011

Dedication

To my parents who are the solid foundation on which I stand

To my dear family who have been my pillar, fountain of love and who have sacrificed immensely throughout this long journey

Declaration

This thesis is the result of my own work. The materials contained in the thesis have not been presented, nor are currently being presented, either wholly or in part for any other degree or other qualification.

Several people contributed in different ways in carrying out the Cochrane review which is summarised in the chapter 3. Joseph Adaktar Lokong independently screened studies for inclusion, assessed risk of bias in the included studies and extracted data from them. Other than supervising the entire process, Professor Paul Garner also independently extracted data from included studies. Sarah Donegan provided advice on statistical methods.

The data set I analysed in chapter 4 was generously provided to me by Heidi Hopkins. The data set came from trial conducted by the Uganda Malaria Surveillance Project (UMSP)—a collaborative initiative comprising researchers at Makerere University in Uganda, the University of California and San Francisco in the US, and the Uganda Ministry of Health (MOH). I did not participate in the design or implementation of this trail. To my knowledge, this data set had, up to the time of this analysis, not been analysed and results presented elsewhere.

Chapter 5 is based on data collected through a health facility assessment conducted by me, assisted by health workers in Gulu and Kisoro districts. Professor Joseph Valadez guided me in developing the survey protocol and tools and mobilised funds which financed all the logistics for the field work (research materials, transport accommodation, and allowances for research assistants).

Acknowledgement

It has been like a journey through a thick dark forest for days on end. I would not have made it through without the support and guidance of many.

First and foremost, my unreserved gratitude goes to my supervisors, Professor Joseph J. Valadez and Professor Paul Garner who patiently walked me through the whole process. This thesis is a reflection of the path you guided me through.

Special recognition goes to my advisory panel members, Professor David Lalloo and Dr. Brian Faragher for their constant encouragement and advice.

Unreserved thanks go to Dr. Heidi Hopkins and colleagues for allowing me to use their data for the analysis in Chapter 4 of this thesis. This was an extraordinary gesture of generosity and trust for which I can not thank you enough.

Sincere thanks go to Uganda Martyrs University for granting me time off work to complete this thesis and for the financial support while on my study leave. Special gratitude also goes to *Doctors with Africa (CUAMM)*—(*dwaCUAMM*)—for the financial contribution towards my accommodation while in the UK. In particular, I would like to sincerely thank the following for making this possible: Dr. John Peter Lochoro, the Country Representative, *dwaCUAMM*, Uganda, and Dr. Stefano Santini, the Project Manager, *dwaCUAMM,* Uganda Martyrs University.

I am most thankful to District Health Officers of Gulu and Kisoro for granting me undue support during the data collection process.

Last but not least, I am thankful to all my siblings and parents for the constant love and support extended to my family during my absence.

Abbreviations and Acronyms

ACT(s)	Artemisinin-based Combination Therapy(ies)
ARI	Acute Respiratory Infection
CBA	Cost Benefit Aanalysis
CEA	Cost Effectiveness Analysis
CHW(s)	Community Health Worker(s)
CI	Confidence Interval
CIPRO	Ciprofloxacin
CMA	Cost Minimisation Analysis
CQ	Chloroquine
CTX	Cotrimoxazole
CUA	Cost Utility Aanalysis
DALY	Disability Adjusted Life Years
ECM	Exchangeable Correlation Matrix
GEE	Generalised Estimating Equations
HC (I, II, III, IV)	Health Centre (referral levels I, II, III, IV)
HFA	Health Facility Assessment
HRP-2	Histidine-rich Protein-2
HW(s)	Health Worker(s)
IMCI	Integrated Management of Childhood Illness
LQAS	Lot Quality Assurance Sampling
MOH	Ministry of Health
NMCP	National Malaria Control Programme
NMFIs	Non-malarial Febrile Illness(s)
NMS	National Medical Stores
OCC	Operating Characteristic Curves
OR	Odds Ratio
ORS	Oral Rehydration Salt
pLDH	Parasite Lactate Dehydrogenase
RA	Research Assistant
RCT	Randomised Controlled Trial
RDT	Rapid Diagnostic Test
R-HFA	Rapid Health Facility Assessment
RR	Risk Ratio, Rate Ratio, Relative Risk
SE	Standard Error
SP	Sulfadoxine-pyrimethamine
STG	Standard Treatment Guideline
UGX	Uganda Shillings
US$	United States Dollar
VHT	Village Health Team
VIT A	Vitamin A
WHO	World Health Organisation
YLD	Years of Life lived with Disability

Abstracts

Introduction
We know that translating new knowledge from research into change in health care delivery is not a simple process. This thesis examines this process for a new technology applied to primary health care in tropical countries: including RDTs in clinical guidelines for treating fever in children

Method
The thesis examines the question: "does implementing policy of using RDTs to target treatment instead of presumptive treatment of fever result in better quality patient care under experimental conditions as well as in routine practice?" Three methodological approaches are used to delineate translation to change in the field. A Cochrane review of randomised trials examines effects on quality of care in a trial, where delivery conditions are usually optimal. An analysis of a dataset from an effectiveness trial from Uganda examines effects of the policy on quality of care delivered within the context of a trial through routine health services. And third, a survey of current practice assesses implementation of an RDT-based guideline when it is introduced into the health system for routine use in selected districts. Across all three components, the thesis examines implementation of the guideline. In addition, both the systematic review and the effectiveness trial measure effects of the intervention on prescribing of antimalarials and antibiotics, and clinical outcomes (primary outcomes).The effectiveness trial evaluates effects of the policy on incremental cost, and the survey of current practice also assesses adequacy of essential health systems inputs and support services.

Results
The systematic review showed that HWs prescribed antimalarials to as many as 40% to 80% of cases with negative RDTs under experimental conditions. Use of RDTs was associated with 29% decline in prescribing of antimalarial drugs. Prescribing of antibiotics did not change in one trial but increased by 19% in another. Data from the effectiveness trial show that HWs used RDTs and adhered to RDT results almost all the time. This reduced antimalarials usage by 60.2% (high), 48.9% (medium) and by 22.1% (low). The data show no significant change in usage of antibiotics. Both the review and the pragmatic trial detected no significant difference in clinical outcomes between RDT and clinical diagnosis arms.

Data from the effectiveness trial shows that use of RDTs is associated with a cost-saving of US$ 0.50 per case of fever (24.5% decline) in low transmission setting, and a cost-saving of US$ 0.33 per case of fever (17.7% decline) in medium transmission. Use of RDTs did not lead to a significant change in cost in high transmission settings: US$ +0.02 (95% CI: US$ -0.97 to US$+1.06). Cost-savings were accrued exclusively in older children and adults.

The survey found inadequate implementation of all components of the guideline in both districts. Essential supplies, equipment and in-service training were inadequate in both districts.

Discussion and conclusion
Antimalarial use is lower when RDTs are used to guide treatment of fever instead of presumptive treatment. This results in savings from drugs costs in older children and adults with fever in low and medium transmission areas. This research does not confirm whether or not use of RDT-based guidelines has any effects on usage of antibiotics or clinical outcomes. A case study of Uganda shows that when delivered through routine services, none of the components of an RDT-based guideline is implemented to acceptable standards. There is insufficient evidence to suggest that the policy is superior to presumptive treatment of fever in terms of clinical outcomes. However, it can save money for medicines in low and medium transmission settings if its use is restricted to older children and adults.

Table of Contents

List of Tables

List of Figures

Chapter 1

Introduction

Chapter 1 Introduction

1.1 Background

1.1.1 Burden of malaria in Africa

Malaria is one of the leading causes of morbidity and mortality in the world. It causes 189 to 327 million clinical cases and more than 800 thousand deaths each year [1]. More than 80% of cases and more than 90% of deaths occur in sub-Saharan Africa [1]. Eighty five percent (85%) of all deaths occur in children under 5 years [1]. Limited evidence from countries in sub-Sahara Africa shows that, in the late 1990's, both poor and better-off households spent approximately $19 for malaria treatment annually [2, 3], and that the annual per capita growth of the GDP in the countries intensely affected by malaria was reduced by 1.3%[1] [5].

In Uganda, malaria accounts for 30% to 50% of outpatient visits at health facilities, 15% to 20% of all hospital admissions, and 9% to 14% of all hospital deaths [1]. In 2006, the reported number of fever episodes suspected of being malaria was 0.94 (0.47 – 1.4) per capita per year, with the prevalence among children younger than 5 years of age being 2.9 episodes (0.54 to 5.5) per child per year [1]. About one third of them , which equates to between 10 and 12 million clinical cases, were treated in the public health sector [3, 6]. Nearly half of hospital in-patient deaths among children less than five years of age were attributed to clinical malaria [3, 6, 7]. In addition, malaria is responsible for up to 22% of low birth weight in newborns in high transmission areas, and a major cause of spontaneous abortions in low endemicity areas [3]. A significant percentage of illnesses and deaths due to malaria occur at home and are not captured by the facility-based Health Management Information System (HMIS). It is estimated that the current morbidity due to malaria reduces economic output in Uganda by 26.3% [8].

1.1.2 Malaria policies in the last 10 years

Prompt diagnosis and effective treatment is the mainstay of the WHO malaria control strategy[1]. However, the diagnosis and treatment of malaria has been problematic. Until recently, WHO guidelines have stressed treating all cases of fever as malaria in endemic

[1] the GDP per capita for Uganda in 1990 was US$ 700 (adjusted for purchasing power parity)4. World Bank/UN Common database, *Globalis, Uganda.*

areas irrespective of the cause of fever[9]. Symptoms-based treatment was intended to simplify the management of malaria in settings without the capacity for microscopy [10]. Although a poor predictor of malaria [11-14], symptoms-based diagnosis has been accepted on the ground that it is better to treat all febrile cases as malaria than to miss one potentially fatal infection due to malaria, especially in a child younger than 5 years old [15]. Further, for several decades, health workers in malaria endemic areas have treated uncomplicated malaria using mono-therapies which were cheap and easy to administer [15].

Owing to the emergence of *P. falciparum* parasites which are resistant to the anti-malarials previously used, the WHO revised the guidelines for treating malaria in 2001, recommending the use of artemisinin-based drugs as the first line drugs for treating uncomplicated malaria [16]. By 2001, most of artemisinin-based drugs were in the form of mono-therapies [16]. In 2006, a new guideline was developed, which recommended the use of Artemisinin-based Combination Therapies (ACTs) for treating uncomplicated malaria instead of artemisinin-based mono-therapies. ACTs have a more complex dosing regimen than all the antimalarials previously used [17]. In addition, ACTs are nearly 10 times more costly than the antimalarials previously used for treating uncomplicated malaria [15, 18].[2] For this reason, the 2006 guidelines also stressed parasitological confirmation of malaria prior to prescribing ACTs, wherever possible, except in children in high transmission areas [9]. In 2010, the malaria policy was revised further, recommending parasitological confirmation in all age groups, including children, and in all transmission settings [19].

Both the 2006 and 2010 malaria treatment guidelines recommend light microscopy as the gold standard diagnostic tool for malaria. Microscopy has several advantages, including parasite quantification, species differentiation and staging of parasites, all of which are useful indicators in selection and evaluation of treatment [20-22]. However, the health system in malaria endemic countries cannot widely deploy microscopes in rural health centres, the first contact points for most febrile patients, in the near future because of the high investment and operational costs, and the technical capacity

[2] For example, in Uganda, the estimated cost of an adult fixed-dose of artemether-lumefantrine is US$ 2.4 in a public facility. The cost of an adult dose of CQ/SP, the combination of antimalarials previously used for treating uncomplicated malaria, is US$ 0.2 in the same type of facility. The cost of the same drugs in the private sector is US$ 10.0 and US$ 1.10 respectively (Obua, 2007)

building required [23, 24]. For these reasons, antigen-based Rapid Diagnostic Tests (RDTs) are proposed for use in settings without capacity for microscopy [9, 19]. Guidelines based on RDTs are now operational in many malaria endemic countries to support the diagnosis and treatment of febrile patients in settings without microscopy [25].

Therefore, over the past 10 years, the policy for treating fever in malaria endemic areas has changed from a symptom-based policy using cheap, easy-to-administer mono-therapies, to parasite-based policies using expensive ACTs with complex dosing regimens. In addition, the current policy introduces a new diagnostic technique into the health system in resource-poor settings with limited supervision. Implementing this policy requires a significant shift in mindset and practice [15, 26].

1.2 Problem statement

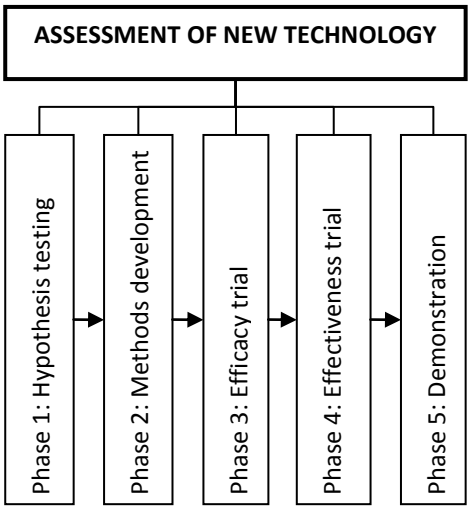
Clinical guidelines are intended to systematically introduce scientific evidence into practice in order to improve practitioners' performance and the quality of patient care [27-29]. Effectiveness and efficiency can only be achieved if there is a change in clinical practice—that is if guidelines are sufficiently implemented [27-29]. Non-adherence to clinical guidelines is common place the world-over [28, 30-39]. In Africa, evidence from routine practice shows wide variability in Health Worker (HW) adherence to parasite-based guidelines for treating patients with fever [32-37]. Presently, it is uncertain if RDT-based guidelines can be implemented sufficiently, and if their use can lead to improved outcomes or cost-savings, especially in routine practice [14, 26, 33, 34]. Further, the health systems capacity to support implementation of the policy has not been demonstrated [26]. It is generally accepted that it is unethical to disseminate and implement an intervention whose benefits are not known for certain [40]. If guidelines are disseminated for wide-scale use without sufficient evaluation, the decision may result in more harm than good being done [26, 40, 41]. Before adopting an RDT-based guideline for routine implementation in settings where the conventional treatment of fever is currently based on clinical judgement, there is need to demonstrate that its application can lead to better quality of patient care.

1.3 Framework for assessing new technologies

New ideas, knowledge, practices, products, guidelines, services are referred to as innovations [41, 42]. The process by which innovations spread and get adopted for routine use is called diffusion [41, 42]. Diffusion of innovations is slow, complex and unpredictable [43, 44]. Several diffusion models based on behaviour change, social change and organisational change theories have identified numerous barriers which must be overcome in order to move innovations from providers to users in routine practice [25, 28, 31, 45-48]. All the models have identified the lack of or inadequate information about the technical and operational attributes of an innovation as a significant barrier to its effective dissemination into the real world [28, 30, 31, 41, 46]. Accordingly, all diffusion of innovation models consider assessment of the value of an innovation and of its applicability in routine practice as a critical step in the diffusion process [41, 47]. Therefore, innovations should be considered ready for application in real-world settings only if they have undergone sufficient evaluation to assess their values and applicability in routine practice [40, 41].

Assessment of new technologies involves several well-recognised level and types of research. Figure 1 depicts the common phases identified by several models which an innovation must pass through during the assessment stage of the diffusion process [27, 40, 41, 49], namely: hypothesis development (phase 1), methods development (phase 2), smaller-scale efficacy trials (phase 3), larger-scale effectiveness trial (phase 4), and demonstration or implementation studies (phase 5).

Figure 1: Phases in assessment of a new technology



Sources: [27, 40, 41]

Phase 1 and phase 2 correspond with the development stage of the innovation where by hypotheses are conceptualised, tested and then transformed into the technology (e.g. devices, protocol, methods, drugs). Assessment of the value of the innovation consists of phase 3 through 5. Phase 3 consists of smaller-scale efficacy trials which use Randomised Controlled Trials (RCTs) to determine whether use of the innovation does more good than harm under optimum conditions [40, 41, 50]. Phase 4 consists of larger-scale effectiveness trials which use pragmatic trials to determine whether use of the innovation does more good than harm when delivered under real-world conditions (e.g. larger population in routine practice) [40, 50].

Phase 5 consists of demonstration or implementation studies to test the quality of delivery of the innovation in real-world settings [40]. It assesses whether all components of the intervention can be delivered to acceptable standards in routine practice. Additionally, demonstration studies may assess operational attributes of the innovation such as ease-of-use, timeliness, acceptability, compatibility and adaptability[40]. Further, demonstrations are useful for identifying potential barriers within the user system that may hinder effective implementation of the innovation in real-world settings—for which interventions may need to be identified and implemented to ensure effective implementation of the guideline [40].

Economic evaluation of the intervention may be undertaken as part of phase 3, phase 4, or phase 5 [47, 51, 52].

Innovations should be considered ready for application in real-world settings only if they have passed through phases 3 through 5 [40, 41]. A new technology may be ineffective in real-world settings either because it is not efficacious or because of poor implementation of an efficacious intervention [40].

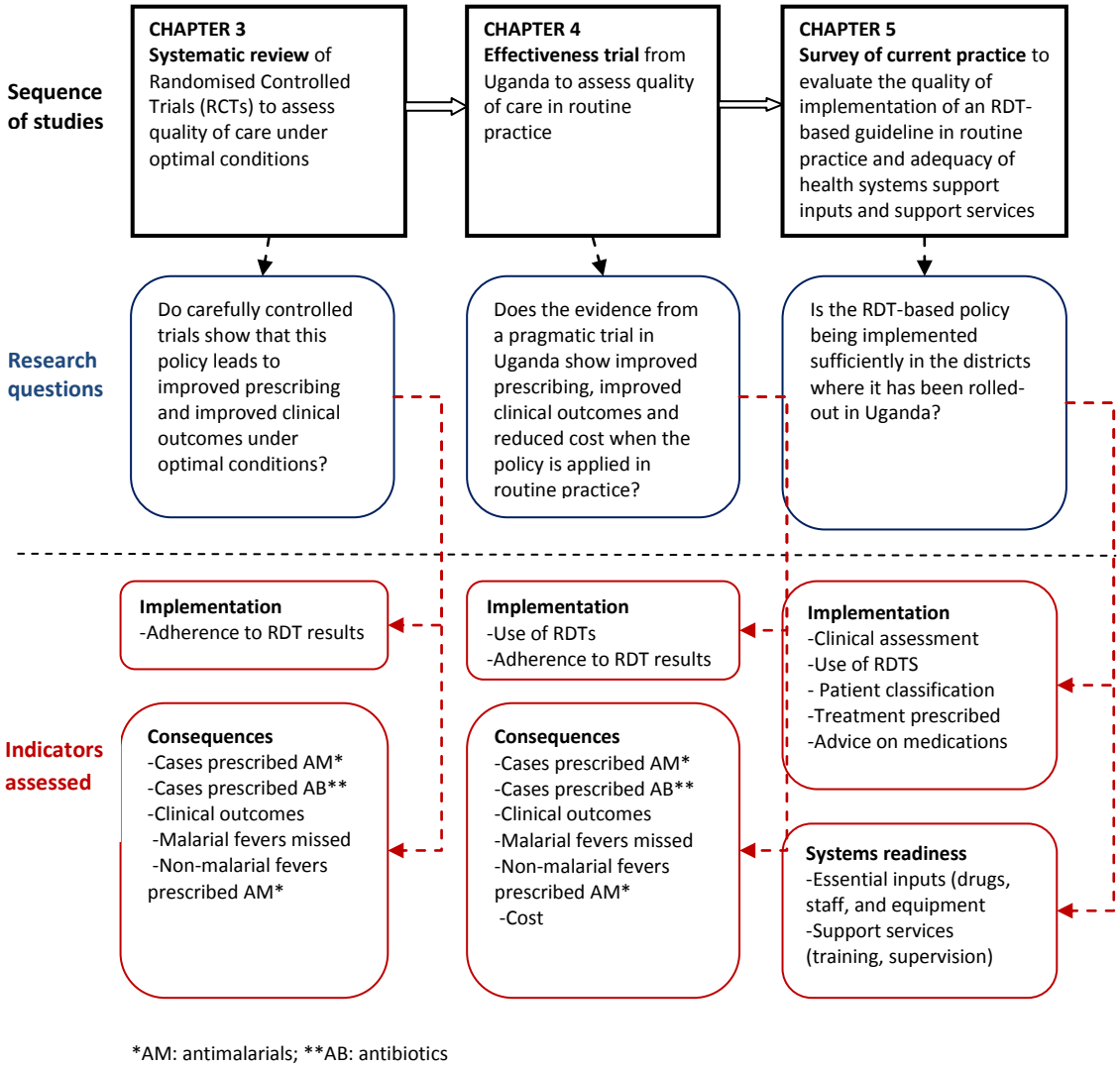
1.4 The thesis

This thesis examines evidence for this evaluation process for a relatively new technology: policy for treating fever, based on the use and results of RDTs in settings without capacity for microscopy. RDT-based policies for treating fever are expected to improve the quality of practice, and thereby reduce usage of antimalarials, lower treatment cost and improve clinical outcomes in settings where presumptive treatment is currently the norm [14, 32-35, 53]. This thesis examines the extent to which these objectives can be achieved when the policy is delivered under carefully controlled conditions and when delivered in routine practice. Further, it assesses the quality of current practice in a sample of districts in Uganda where an RDT-based guideline has been rolled-out for small-scale implementation. This assessment is carried out to determine whether all components of the guideline can be delivered to acceptable standards in routine practice where there is relatively limited supervision and resources compared to implementation in a research setting; and whether essential health system inputs and support services which are required for effective implementation of the guideline are adequate.

1.5 Conceptual map of the thesis

The thesis consists of 3 studies which reflect the sequence of movement from phase 3 (efficacy trial) to phase 5 (demonstration) of the assessment process (Figure 2).

Figure 2: Conceptual map of the thesis



*AM: antimalarials; **AB: antibiotics

A systematic review of evidence from randomised trials in Africa is undertaken to evaluate the effect of RDT-based policies on the amount of antimalarials and antibiotics prescribed and on clinical outcomes under relatively optimal conditions. In addition, the review assesses the extent to which HWs prescribe antimalarials according to RDT results, which is an indicator of HW adherence to RDT-based guidelines. Other aspects of guideline implementation, such as the quality of clinical assessment, use of RDTs,

patient classification (diagnosis given by the HW) and advice on medications are not evaluated in this review due to data limitation.

An analysis of existing dataset from a pragmatic trial from Uganda is used to evaluate the effect of RDT-based policies on the amount of antimalarials and antibiotics prescribed. Additionally, the analysis examines the effect of the policy on clinical outcomes and incremental cost in routine practice in Uganda. Further, it describes HW's use of RDTs and response to RDT results in routine practice. Because of data limitations, it does not assess the quality of implementation of other components of the guideline such as clinical assessment, patient classification and advice on medications.

A survey of current practice was carried out in Uganda to evaluate the quality of actual practice and adequacy of inputs and support services in the districts where RDT-based guidelines have been rolled out for small-scale implementation.

1.6 Research question

Does implementing an RDT-based guideline instead of presumptive treatment of fever result in better quality of patient care under optimal conditions as well as in routine practice?

1.7 Aim

To establish if the use of RDT-supported guidelines in treating fever leads to better quality of patient care relative to presumptive treatment when delivered under optimal conditions and through routine clinical practice.

1.8 Objectives

1) To review evidence from RCTs to determine if RDT-based policies for fever lead to better quality of patient care than presumptive treatment under optimal conditions

2) To analyse data from a pragmatic trial to determine if RDT-based policies for fever lead to better quality of patient care than presumptive treatment in routine practice

3) To assess the quality of actual practice when an RDT-based guideline is rolled out to health services in a district as a whole.

1.9 Overview of the thesis

The thesis is structured into 6 chapters.

Chapter 1 Introduction

Chapter 1 describes the evolution of the policies for treating fever in malaria endemic countries over the past 10 year, including the reasons for the shift to parasite-based policies. It outlines the unanswered questions regarding the utility of RDT-based policies in routine practice, which this thesis attempts to answer. In addition, it describes an evaluation framework which guided the structure of this thesis and the selection of variables assessed. It provides a conceptual map showing the different studies undertaken by the thesis, outlining the questions each study attempted to answer and the common variables examined across studies as well as the specific ones examined by each study. Further, the chapter gives a brief description of the Uganda health system including an outline of the RDT-based guideline for Uganda.

Chapter 2 Literature review

This chapter summarises the theories and models which describes the process of introducing a new technology into a user system, and which offers a comprehensive framework for evaluating a new technology. It describes relevant concepts and methods used in evaluating health technologies or quality of care and it justifies the methods used in the thesis. It highlights the technical and operational attributes of different types of malaria RDTs. In addition, it gives an account of the performance of various types of RDTs in different epidemiological contexts. Further, it reviews evidence from Africa on the quality of guideline implementation and effectiveness of interventions aimed at improving the quality of guideline implementation. Finally, it describes the principles of the Lot Quality Assurance Sampling (LQAS) method and how the method can be applied in healthcare to monitor the quality of health care.

Chapter 3 A systematic review

Chapter 3 consists of a systematic review of RCTs from Africa which have examined effects of RDT-supported treatment of fever relative to presumptive treatment on prescribing and clinical outcomes. Further, it shows the extent to which HWs prescribed antimalarials according to results of RDTs under carefully controlled conditions. The findings of this review show the quality of practice and "efficacy" of using RDT-based guidelines under relatively optimal conditions.

Chapter 4 Analysis of data from an effectiveness trial

Chapter 4 analyses an existing set of data from an effectiveness trial from Uganda to determine the effects of using RDT-supported guidelines in routine practice. It assesses whether introducing an RDT-based guideline for fever leads to improvement in prescribing, clinical outcomes and healthcare cost. Additionally, it assesses whether HWs requested RDTs and responded to RDT results as expected. Further, the analysis evaluates how local malaria prevalence and age profile of the target population all combine to impact the quality of prescribing, clinical outcomes and incremental cost in routine practice. The findings of this analysis show the quality of practice, effectiveness and cost implications of implementing an RDT-based guideline in routine practice.

Chapter 5 Survey of current practice

Chapter 5 consists of a survey of current practice to evaluate the quality of actual practice and adequacy of support services in the districts where RDT-based guidelines have been rolled out for small-scale implementation. The study uses Health Facility Assessment (HFA) tools based on the Lot Quality Assurance Sampling (LQAS) method, which provides a rapid comprehensive assessment of key diagnosis and treatment subsystems stipulated in the guideline. In addition, the HFA tools allow for an assessment of key health systems inputs and support activities (e.g. training, supervision). The survey uses the LQAS method to classify implementation of the guideline in the selected districts as adequate or inadequate, basing on a pre-defined performance benchmark.

Chapter 6 General discussion

Chapter 6 discusses the main findings in Chapters 3 to 5, comparing and contrasting outcomes which were assessed across studies.

Chapter 2

Literature Review

Chapter 2 Literature Review

2.1 Introduction

This chapter is organised into eleven sections. Section 2.3 describes the process of knowledge translation based on diffusion of innovation theory and the implication of the diffusion theory for the transfer of RDT-based policies into the health system. Section 2.4 defines the focus of the thesis and justifies the choice of a 3 chapter approach. Section 2.5 consists of definitions of relevant concepts and methods which are used in evaluation of new technologies and quality of care; and which have been applied in this thesis. Section 2.6 describes the basic principle of Lot Quality Assurance Sampling (LQAS) method and its application in healthcare for monitoring interventions in developing countries. LQAS method is used in Chapter 5 to assess the adequacy of adherence to RDT-based guidelines. Section 2.7, provides an overview of several biologic malaria diagnostics that have emerged over the last 100 years. Section 2.8 provides an overview of malaria RDT technology. It outlines the factors affecting performance of RDTs, an overview of the WHO product testing initiative, and performance of various RDT assays under experimental conditions, in different epidemiological settings. Section 2.9 consists of a synthesis of evidence from Africa on the quality of implementation of various clinical guidelines. Section 2.10 synthesises evidence from Africa on the effectiveness of various interventions aimed at improving the quality of implementation of various clinical guidelines in Africa. Section 2.11 consists of a summary of the literature review, which brings together all the main topics discussed in the chapter, and describes their relevance to the questions the thesis attempts to answer, methods used in the thesis, and how they have informed the discussions in chapters 3, 4, 5 and 6.

2.2 Search strategy

Searches of electronic databases were conducted mainly through EBSCO interface hosted at the University of Liverpool Library. Searches were conducted in the following databases: Medline, CINHAL, PsychINFO, and the Cochrane Database for Systematic Reviews. Search terms were applied using Boolean operators as shown in Appendix 3. Searches were restricted to articles published between January 1990 and December 2011. The hits retrieved were refined using MeSH (Medical Subject Headings) terms. Titles of articles retrieved were scanned, and abstracts of the articles considered relevant

to the objectives of the thesis were reviewed. Full texts of relevant articles were retrieved from the University of Liverpool Library database. Articles which were not available electronically were obtained through the University of Liverpool Library Holdings. The results of the searches, the papers retrieved and papers reviewed are summarised in Table 34 below.

Table 1: Number of articles retrieved and reviewed from electronic databases

Section		Number of hits retrieved after applying MeSH terms		Number of full texts retrieved	Number reviewed
		Before removing duplicates	After removing duplicates		
2.3	Theory of diffusion of innovations	236	124	25	22
2.5	Concepts and methods	27	27	27	23
2.6	LQAS technique	1324	115	47	13
2.7, 2.8	Overview of malaria diagnostics and malaria rapid diagnostic tests	1324	835	25	18
2.9	Implementation of clinical guidelines in Africa	392	297	20	15
2.10	Effects of intervention on quality of implementation of guidelines (Africa)	392	297	20	9

Citations of electronically selected articles were downloaded and managed using EndNote X4. Searches were performed throughout the thesis writing and citation library updated regularly.

Additional searches were conducted from the websites of the WHO and the Ministry of Health (Uganda)[3]. Further, I checked the reference lists of all the selected full texts, and searched for relevant articles or reports in the databases and websites described above or in Google scholar.

[3] The website of the Uganda Ministry of Health was search for policy documents and/or guidelines

2.3 Theory of diffusion of innovations

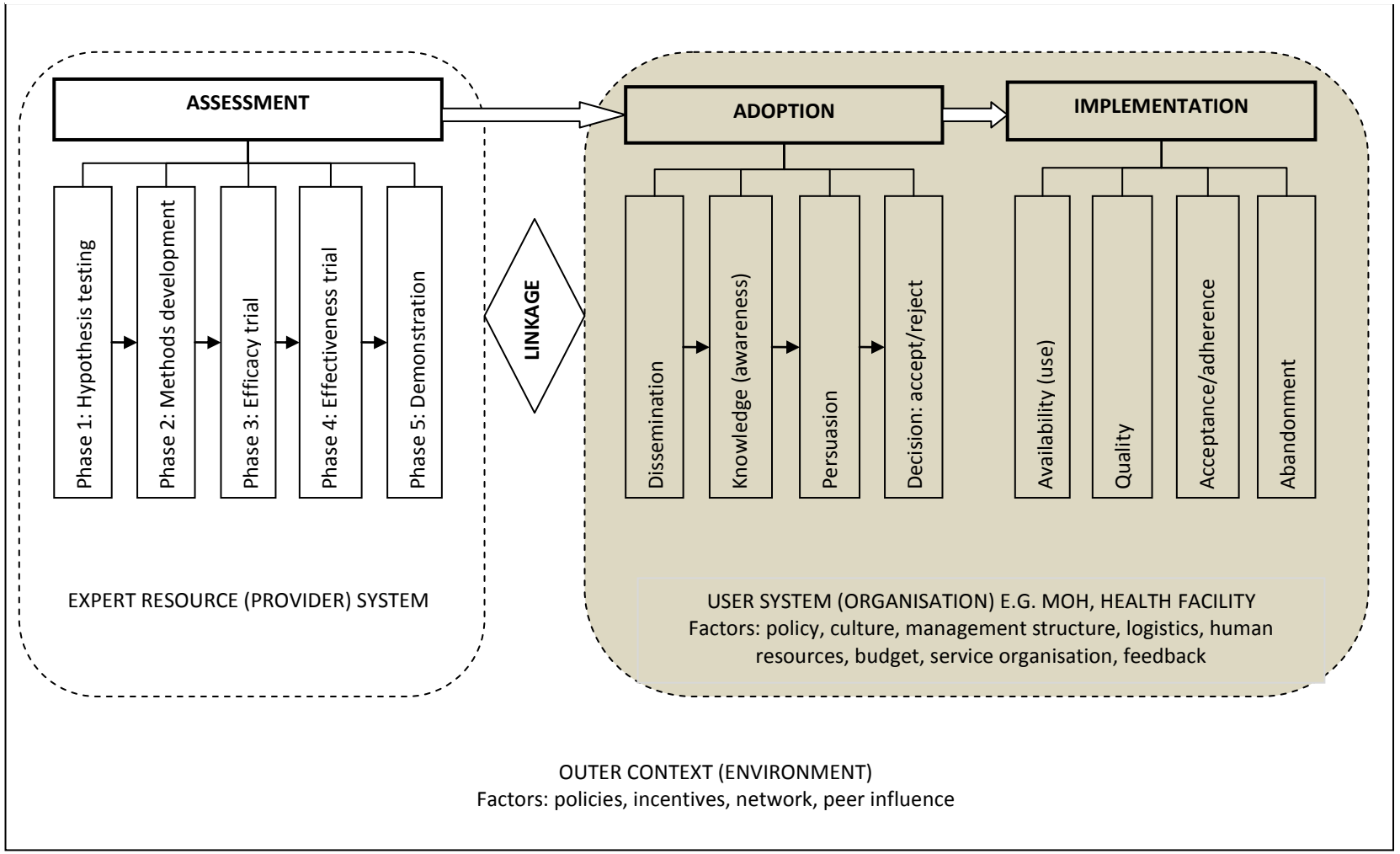
2.3.1 Classical diffusion model

According to classical diffusion models, innovations originate from some expert resource system (e.g. research organisation), which then diffuses it to potential users. Traditionally, the process consists of a series of steps which can be grouped into 3 main phases (Figure 3): (a) development and assessment, (b) adoption, and (c) utilisation (routine implementation) [47].

At least five types of assessment may be required before an innovation is considered ready for adoption and wide-scale use: assessment of safety, assessment of efficacy, assessment of effectiveness, assessment of cost, and assessment of applicability in real-world settings [41, 42]. Upon completion of assessment of the attributes of the innovation and its applicability in the real-world, it is then deemed ready for introduction into the user system [40, 41]. Dissemination into the user system is undertaken either by the developer or an external agent [41, 46, 54].

Introduction of an innovation into a user system sparks off an adoption process. Adoption refers to the decision to accept or reject an innovation [28] or whether to accept it now or to defer the decision to a later date [47]. Adoption decision may be taken at a central (policy level) and/or by individual users [47]. A vital component of the adoption process consists of dissemination of innovations into the user system, by means of which potential adopters and users are made aware of availability and attributes of innovations, and by means of which they may be persuaded to adopt them [31, 46].

Figure 3: Diffusion of innovation model



Implementation consists of usage activities that follow the adoption decision [46]. More specifically implementation refers to the extent to which the technology is made *available* to the target audience in a manner that is *acceptable* to them [40]. Therefore, implementation studies assess whether all components of the technology are being delivered to the target audience to acceptable standards, and the extent to which the target audience are receptive to, participant in, comply with or adhere to all the components [40, 47]. Additionally, implementation studies seek to identify barriers to adoption and implementation of innovations and to test the effectiveness of interventions intended to address them [29, 43, 54]. Implementation studies may be undertaken as part of technology assessment (phase 5) [40, 41].

The classical diffusion model just described depicts a centralised, hierarchical (top-down) model whereby innovations are transferred from the provider system to the users system as a matter of policy [41]. Users have no control over which innovations are considered. Centralised models depict what typically happens in a government system [41]. Diffusion of innovations in the private sector is generally decentralised: innovations arise from multiple sources and users exercise control over which innovations to adopt [41, 46]. Innovation whose diffusion follows a centralised model (e.g. the policy under evaluation) are more likely to be adopted than those for which adoption decisions are decentralised [41].

The classical diffusion model also portrays a linear transfer pathway of innovations from development to implementation. In practice, the process may be iterative or cyclical and some steps may be skipped. For example, innovations may be adopted without prior assessment or assessment may be undertaken after adoption [40].

2.3.2 Drivers of diffusion of innovation

Each step in the diffusion process is associated with barriers which must be overcome in order to move the innovation from assessment stage (provider system) to routine use in the real world (user system) [41]. Several models based on behaviour change, social change organisational change theories have identified 6 main levels and types of factors which might affect the adoption and implementation of innovations. They are summarised in Table 2 below.

Table 2: Factors which determine the rate of diffusion of innovations

System	Factors	Examples
Provider system	Attributes of innovation	Relative advantage, simplicity, and compatibility
User system	Attributes of individual professional (user)	Knowledge, attitude, motivation to change, behavioural routine
	Attributes of patients	Expectations, knowledge, attitude, compliance
	Social context	Opinion of colleagues, culture of the network, collaboration, leadership
	Organisational context	Organisation of care processes, staff, capacities, resources, structures
	Economic and political influences	Financial arrangements, regulations, policies

Adopted from Grol and Wensing[38]

a) Attributes of the innovation

Attributes of an innovation have repeatedly been identified as a key determinant of whether or not, or how fast it is adopted for use in routine practice [28, 30, 31, 38, 41, 46]. Attributes of innovations are claimed to impact on the ability to effectively disseminate them into the user system [41]. Therefore, assessing attributes of an innovation is essential, not only for ethical reasons, but also because it has implications for effective dissemination of the innovation. Attributes of an innovation may be technical or operational in nature, all of which need to be assessed prior to dissemination [20-22]. Technical characteristics include properties such as efficacy, effectiveness, efficiency and safety. Operational characteristics consist of features that affect the application of innovations, e.g. simplicity (ease-of-use), timeliness, compatibility and adaptability [28, 30, 31, 38, 41, 46].

Technical attributes of an innovation appear to have the most powerful influence on the behaviour of potential users and policy makers [28, 30, 31, 41, 46]. Innovations that are perceived to be more beneficial than the methods previously or currently in use are more likely to be adopted [41, 55]. Therefore, the more uncertainty is reduced by providing sufficient evidence on the efficacy, effectiveness, cost and safety of innovations, the more likely that potential users will adopt them [41, 55].

Simple innovations are understood and therefore diffuse more rapidly than complex ones [28, 31, 42, 55]. Complex innovations consists of many different components [30]. An RDT-based guideline is a typical example of a complex innovation as it consists of many inter-related components and its diffusion and effective implementation could be problematic [41, 46].

Further, an innovation is more likely to be adopted if its use is compatible with the norms, values, and perceived needs of the potential users, patients and policy makers [28, 31, 42, 55]. An innovation which requires change in existing routine and habits is less likely to be adopted. In addition, an innovation which invokes negative reactions in patients or other key players because it does not fit in their expectations is less likely to be adopted [28, 31, 42, 55].

b) Attribute of the individual professional

Knowledge

According to behaviour change theories, individual health workers are more likely to accept and use a new technology if they have good understanding of the aims, content, scientific merit and how to use it. An innovation is less likely to be adopted if its use demands the acquisition of new competence (skills). Effective communication of the technical characteristics of an innovation is a prerequisite for change in attitude and practice [28, 31, 46].

Importance of training and experience

According to human capital theory, the knowledge required to effectively use an innovation is accumulated through pre-service and in-service training (induced knowledge) [46, 56], or though learning-by-doing (experience) [28, 31, 42, 55, 57].

Experiential knowledge is accumulated by performing the same tasks repetitively and consistently [57-59]. The knowledge acquired from a previous experience is carried forward to the next task [59] resulting in sequential and incremental accumulation of knowledge [57-59]. Thus, the stock of knowledge accumulated through learning-by-doing is proportional to the volume of the task performed [56-59].

Experiential knowledge has greater impact on the health worker's ability to correctly use or apply a new technology than induced knowledge [56]. Consistent practice serves to 'top up' the knowledge base acquired through training, until an optimum or maximum level is reached. Experiential knowledge usually accounts for the bulk of the knowledge stock accumulated by a user and

consistent practice serves to maintain an optimum knowledge base and clinical proficiency over time [56]. On the other hand, lack of or inadequate practice results in deterioration in knowledge and skills over time (forgetting). The rate of forgetting is directly related to the length of interruption in practice [59, 60]. Therefore, users of a given technology need to perform a minimum volume of tasks with the technology in order to maintain an optimal level of proficiency [59]. Furthermore, experience is thought to result in the development of more finely developed skills, which in turn leads to better performance—the notion of practice makes perfect—for example being able to execute relevant procedures more appropriately and more quickly [56, 61].

Although regular refresher courses or continuing medical education can reduce the rate of forgetting [59], their effects are short-term [62]. Consistent use of a technology is essential in maintaining an optimum level of knowledge and clinical proficiency over time [59].

c) Peer influence and leadership

Social change theory suggests that adoption and correct use of an innovation can be brought about through peer influence and through leadership. Peer influence is exercised through individual interactions or professional networks, while leadership normally takes the form of advice from senior and respected professionals on task issues (e.g. support supervision) [38, 46].

d) Patients' expectations

Patients' expectations influence prescribing practices. Clinical recommendations that do not fit patients' expectations may invoke a conflict of interest between patients and doctors [30]. The evidence on this theory is equivocal [63] and few studies have included it in the models for analysing the determinants of adherence to guidelines [38, 64].

e) Organisational factors

Organisational change theories suggest that, irrespective of the inclination of individual clinicians, the process of change in practice can be facilitated or impeded by organisational factors; such as staffing, financial resources, availability of required inputs, organisation of activities, and organisational policies, [28, 31, 38, 41, 46, 48, 55].

f) External policy environment

According to classical diffusion models, innovations whose diffusion follows hierarchical (top-down) pathways (e.g. the RDT guideline) are more likely to be adopted than when adoption decision is decentralised [41].

2.3.3 Implications for introducing RDT-based policies

The diffusion model depicts the type and sequence of research required in assessing the value and applicability of a new healthcare technology before dissemination for routine use. Phase 1 and phase 2 concern product developments. In the case of the policy under evaluation, this process includes evaluation of the intrinsic properties of various RDTs through laboratory-based product testing and field trials: diagnostic accuracy, thermal stability and ease-of use. These assessments aim to drive the quality of the products to an acceptable level before they are recommended for use in clinical settings[65]. An overview of RDT field trials and WHO product testing and is provided in sections 2.8.3 and 2.8.5 respectively.

This thesis is concerned with phase 3 through 5 of technology assessment process. Phase 3 and phase 4 studies provide information on efficacy, effectiveness and efficiency. Phase 5 study assesses whether the guideline can be implemented sufficiently in routine practice. Additionally, it identifies key factors within the user system which can affect implementation of the policy in routine practice—which may need to be investigated, or which may need to be addressed prior to or during wide-scale implementation. These may include the knowledge and skills required to use RDTs and to apply all the components of the guideline; and effective communication strategies to impart the required knowledge. It also implies that is vital to maintain adequate supply of essential logistics (RDTs, drugs, equipment). Otherwise, inconsistent application of the policy could lead to forgetting and deterioration in performance. Technical support from senior professionals is equally necessary in facilitating the process of change.

2.4 Choice of a 3 chapter approach

The diffusion of innovation model outlines the sequence and type of research required in assessing a new technology before dissemination for routine use. Phases 1 and 2 correspond to the development phase of the technology and involve evaluating the intrinsic properties of the technology [40, 41, 50]. Phases 3 and 5 involve evaluating the usefulness of the technology while phase 5 assesses its potential applicability in routine practice.

Since their introduction in the 1990s, malaria RDTs have been characterised by inconsistency in manufacturing standards, product modification, withdrawal, insufficient quality control and variability in product stability [24, 66, 67].

Several field evaluations have been carried out to assess assay performance characteristics, especially diagnostic accuracy. Limited laboratory-based evaluations of malaria RDTs have also been conducted by manufacturers by testing assays against panels of blood infected with malaria parasites [20, 22, 65]. However, comparability of the results of these studies has been limited by errors associated with trial designs, variation in reference standards, insufficient quality control and epidemiology of malaria [20, 22, 65].

Because of these limitations, guidelines have been introduced to assist diagnostic assay development and manufacture, regulatory approval processes, and to support malaria control programmes [65, 68, 69]. Additionally, WHO has introduced product testing imitative that assesses assay performance characteristics in a standardised way. The WHO product-testing initiative has been running since 2006. Three rounds of product-testing have been completed, and a 4th round is on-going. These assessments aim to drive the quality of RDT products to an acceptable level before they can be recommended for use in clinical settings[65]. An overview of RDT field trials and WHO product testing is provided in sections 2.8.3 and 2.8.5 respectively.

Therefore, extensive phase 1 and phase 2-level evaluations have already taken place with regard to RDTs. The major debate regarding RDTs currently is whether using them in clinical settings to guide treatment in fever patients can lead to improved quality of healthcare. It is presumed that RDT-supported management of fever could avert irrational use of ACTs, thereby resulting in cost-saving [14, 32-35, 53]; improve diagnosis and treatment of parasite-negative individuals [19, 70-72]; limit the risk of resistance and adverse reactions [72]; and thus lead to better clinical outcomes in

fever patients, compared to presumptive treatment. Before wide-scale implementation, there is need to demonstrate that RDT-based guidelines can be implemented sufficiently and that implementing them instead of guidelines based on clinical diagnosis can lead to better outcomes [14, 26, 33, 34].

According to diffusion of innovation model, in order to determine if a new technology is effective in routine practice, it is important to know if it is efficacious first.

Accordingly, the first level of research undertaken consists of an assessment of the efficacy of the intervention. Assessment of efficacy is considered a necessary step in the development and diffusion of a new technology or intervention [40]. Knowledge of efficacy is necessary to inform the decision as to whether or not it is necessary to carry out an effectiveness study and to aid interpretation of evidence from effectiveness trials. Once an intervention is shown to be efficacious, it is then useful to carry out an effectiveness trial to investigate if it can work in real-life settings; and implementation research to assess if all components can be delivered to acceptable standards in routine practice in a manner that is acceptable to the target audience [40].

The second level of study chosen is an effectiveness study, which assesses if the intervention works for the general population. The third level of research chosen is a demonstration or implementation study. Demonstration studies can show the extent to which various elements of a multi-component intervention are actually implemented [73, 74]. In addition, it enables the identification of facilitating and inhibiting contextual factor.

2.5 Concepts and methods

2.5.1. Malaria endemicity and classification

Malaria transmission levels vary significantly between different geographic regions [75-77]. Grading of malaria transmission is important for targeting control interventions.

Malaria endemicity is typically classified into 4 levels, based on a variety of quantification methods [77, 78].

(a) Holoendemic, which is characterized by very intense, year-round transmission, resulting in severe anemia during early childhood but considerable degree of immunity outside early childhood.
(b) Hyperendemic, characterized by intense but seasonal transmission. Immunity is insufficient in all age groups, and cerebral malaria is common in older children.
(c) Mesoendemic, characterized by regular seasonal transmission under normal rainfall conditions. Transmission is low in dry seasons. Cerebral malaria is a common feature. Mesoendemicity is typical in communities in subtropical zones.
(d) Hypoendemic, where transmission is low and intermittent and the effect on the general population is unimportant. However, outbreaks of severe malaria and mortality are common in both children and adults.

Four main methods for grading the magnitude of malaria in the population are described in the literature, none of which appears completely satisfactory. They consist of (a) 2 host-based methods which involve estimating the prevalence of palpable splenomegaly (spleen rates) and/or parasitaemia (parasite rates) in the population [77-79], (b) a vector-based method which estimates malaria prevalence in terms of transmission intensity [77, 79], and (c) a vector-based mathematical model which measures endemicity in terms of stability of transmission[77-79]. The descriptions and the relationship between these indices are shown in Table 3 below.

Table 3: Criteria for classifying malaria endemicity

Criterion	Endemicity				Source
	Hypoendemic	Mesoendemic	Hyperendemic	Holoendemic	
Description	Low	Moderate	High	High	[77]
Spleen rate in children 2 – 9 years	0 – 10%	11 – 50%	>51 – 75% Also high in adults (> 25%)	>75% Low in adults	[77, 78]
Parasite rate in children 2 – 9 years	0 – 10%	11 – 50%	>51 – 75%	>75% (among infants aged 0 – 11 months)	[77, 78]
Annual Entomological Inoculation Rate (AEIR)	<0.25	0.25 – 10	11 – 140	>140	[77]
Stability	Unstable	Intermediate	Stable	Stable	[79]

Source: Adapted from Mendis (2009), Hay (2008) and Roll Back Malaria (undated) [77-79]

a) Spleen rate

Spleen rate or spleen index is the method used in most malaria metric surveys to define malaria endemicity[77-79]. It is defined as the proportion of a selected age-group of the population with palpable enlargement of the spleen. It is measured per 100 individuals of similar ages; typically in children aged 2 to 5 years [77, 78]. This method was the first used to quantify malaria disease in the population, having been introduced in India in 1948 [79]. It is determined through a survey of a selected age-group of randomly sampled population. The quantity measured is a point prevalence of splenomegaly, although the term "rate" is often used [79].

b) Parasite rate

Parasite rate is another method that has been used in malaria metric surveys because of its higher specifity than the spleen index [77]. It is defined as the proportion of the population of similar ages with parasitamia [77, 78]. It is also measured per 100 individuals of similar ages; typically in children aged 2 to 5 years [77, 78]. It is determined through a survey of a selected age-group of randomly sampled population. It involves assessing the presence of asexual malaria parasite in peripheral blood by slide microscopy [79]. It represents the point prevalence of parasitaemia in the selected group of the population [78]. Its usefulness in malaria epidemiology remains questionable [79].

c) Transmission intensity

This measure depends on the capacity of the vector (*Anopheles* mosquito) to transmit malaria parasite during its life time. The measure of transmission intensity is Entomological Inoculation Rate (EIR), which refers to the average number of infective bites by *Anopheles* mosquito per person per unit time, usually one year (i.e. Annual EIR = AEIR). EIR is normally measured through sentinel surveillance, using various methods [77, 79].

d) Stability of transmission

This is a mathematical model based on EIR. The classical mathematical model is the Ross-Macdonald model which can be used to predict the relationship between endemicity as measured by *P falciparum* parasite rate (*Pf*PR) and transmission intensity measured by *P falciparum* EIR (*Pf*EIR) [79]. The Ross-Macdonald model shows that *Pf*PR is very sensitive to small changes in *Pf*EIR at low transmission intensity, but it is

insensitive to small changes in *Pf*EIR at high transmission intensity [79]. The Ross-Macdonald model classifies malaria endemicity into stable malaria and unstable malaria [79]. Stable malaria implies an overall balanced or constant presence of malaria, with persistently high prevalence of infection, which is insensitive to environmental changes. Transmission is year-round although there may be seasonal fluctuation. Unstable malaria implies irregular transmission of malaria in space and time. The background immunity in the population is low and therefore the risk of malaria epidemic is high [78]. Intermediate stability is designated between these two extremes. Thus unstable malaria settings consist of hypoendemic areas, intermediate stability corresponds to meso-endemic areas, while stable malaria settings comprise hyper and holoendemic areas.

e) Classification method adopted in this thesis

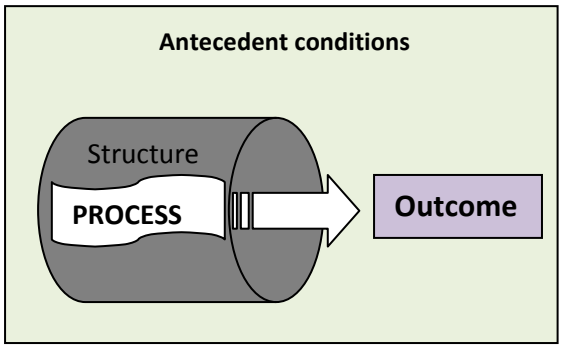
In this thesis, malaria prevalence in Uganda is described as low, medium and high. This classification has been used in previous studies, including the trial analysed in chapter 4 [80, 81]. It is based on AEIR indices derived from a recent 1-year long entomological surveillance data obtained from 7 ecologically different sentinel sites throughout the country, including the study locations mentioned in this thesis [75]. The values of EIR and endemicity classification attributed to the locations studied as part of this thesis correspond to the grading presented in Table 3.

2.5.2. Measuring quality of care

a) Definitions

Definitions of quality vary widely, and so are approaches to measurement [82]. A unifying conceptual framework for assessing the quality of healthcare is the Donabedian model which was developed over 30 years ago [83, 84]. The Donabedian model describes a systemic approach to assessing quality of care. It considers quality on the basis of structure, process and outcomes [84, 85]. Figure 4 below shows the causal relationship between structure, process and outcomes sub-systems.

Figure 4: The Donabedian model of health system performance



Adapted from Donabedian (1980) and Coyle 1999 [83, 84]

Structure describes the setting in which healthcare is provided [83, 85]. It consists of institutional factors such as inputs (human, financial, equipment), policies and regulations, amenities, and provider characteristics (ownership) [83, 85, 86]. Structural attributes measure the presumed capacity of the provider to offer quality health care [86].

Process consists of components of an encounter between a HW and a patient, or what is done to patients (e.g. consultations and laboratory tests) [83, 85, 86]. Measurement of quality in terms of process data is today commonly referred to as performance measurement [86]. The Donebanian model distinguishes quality of process into technical quality and inter-personal (service) quality [83, 87]. Technical quality in healthcare is the extent to which clinical aspects of care meets pre-defined standards of acceptable or good care [87]. It measures disease-oriented aspects of care and deals with what the patient receives relative to what is known to be effective [86, 88]. Interpersonal (service) quality includes aspects such as patient advice, answering questions from patients and taking into account patients' preferences in decision-making [83, 86, 89].

Outcomes attributes consist of clinical endpoints (laboratory values, morbidity and mortality), functional status (physical, mental, social), general well-being (perception, vitality, fatique) and satisfaction with healthcare [83, 85, 86].

The structure-process-outcome model does not operate in a vacuum, but in an environment referred to as antecedents. Antecedents are factors that affect the structure, process and outcomes of medical care, but may not be within the influence of the health provider. Antecedents consist of the environmental context of individuals (policies, culture, beliefs, sanitation) and individual personal characteristics (genetics, socio-demographics). Antecedent factors are known to have the strongest influence on outcomes [83]. Analysis of antecedent factors can help explain whether outcomes are due to interventions or patient factors or environmental factors [83].

For quality of care derived from structural data to be valid, variation in attributes of structure assessed should lead to differences in process and outcomes quality indicators. Further, for quality of care derived from process data to be valid, variations in the process of care attributes measured must lead to a change in outcomes that are important to patients. Similarly, outcomes measures are valid indicators of quality only to the extent that differences in outcomes can be linked to changes in the process of care [83, 86]. Linking outcomes data to process data can inform health managers of the kind of actions they should take when health outcomes are poor [86].

The relationship between structural standards and either process or outcomes attributes is questionable [86]. It is acknowledged that compliance with structural standards does not mean that high quality care is being provided; nor does their use in quality assessment imply that high quality care cannot be provided unless these standards are complied with[86]. Nevertheless, structural standards are often combined with measures of performance and outcomes in assessing quality of healthcare [86, 90-93]. On the other hand, process data are widely considered to be a sensitive predictor of outcomes [94-96]. Therefore, both process and outcomes data are considered to provide valid information about the overall quality of care. However, outcomes data are usually less sensitive in predicting quality of care. This is due to the fact that differences in outcomes may not be related to factors under the control of the HW. For example, patients receiving the same treatment may experience different outcomes due to factors

such as differences in personal characteristics (antecedents). Therefore, a good outcome does not imply adequate quality of the process of care, and vice versa [85].

b) Methods of assessing technical quality of care

Methods of assessing quality of care may either be implicit or explicit. In implicit methods, standards for judging quality of care are not defined a priori. Therefore, there is no basis for judging adequacy of quality. On the other hand, explicit approaches involve the use of standards which are set a priori, on the basis of which quality can be judged as either adequate or inadequate [85]. The assessor then compares what was done against what should have been done and the result is expressed as the proportion of the criteria that were met. Additionally, or alternatively, various levels of quality of care which have been defined on the basis of explicit a priori criteria can be used to predict future outcomes by means of valid models [85].

Results of quality assessments vary with the methods used. Explicit process methods are the most strict, while implicit outcomes-based methods are the least strict. For example, explicit process indicators may show that two population groups have received care that is starkly different in quality. However, implicit (subjective) outcomes indicators of the same care may show a small difference in quality between the two groups. For this reason, it has been proposed that action-oriented assessments of quality of care should focus on process data rather than on outcomes data, especially in assessment of HW performance [85].

c) Implications for RDT-based guideline

The RDT-based policy is aimed at improving the quality of the process of care for patients with fever. Improvement in outcomes is expected to accrue as a consequence of improvement in the quality of practice. Therefore, change in the quality of clinical practice is a more direct outcome of the intervention rather than change in clinical outcomes. Further, it can be construed from the foregoing that change in clinical outcomes may not be a sensitive indicator of whether improvement in the overall quality of care has taken place. On the other hand, indicators of HW performance may be a more valid measure of whether introduction of the policy has led to a change in the overall quality of care. However, we know that the main objective of introducing the new policy is an expected improvement in outcomes and efficiency (cost-saving). Therefore, while indicators of HW performance may be used to demonstrate if change

in the overall quality of care has taken place, it is also vital to show that implementing the new policy does not do more harm than good, relative to the existing symptoms-based guidelines. Ultimately, adoption decisions are supposed to be based on outcomes indicators. However, lack of improvement in clinical outcomes should not be considered to reflect a lack of improvement in the overall quality of care.

2.5.3. Difference between efficacy and effectiveness trials

Table 4 outlines the differences between efficacy and effectiveness trials which, respectively, correspond to phase 3 and phase of technology assessment.

Table 4: Characteristics of efficacy and effectiveness trials

CRITERIA	EFFICACY TRIAL	EFFECTIVENESS TRIAL
Question	Can it work in ideal or optimal conditions	Does it work under routine or "real life" conditions?
Purpose	Explanatory: what works? How does it work	Pragmatic: should the service be provided to a wide variety of the population in a wide variety of settings?
Participants	Smaller scale More homogenous; typically includes subgroup analysis	Larger scale More heterogeneous; reflective of the general population or real life. Subgroup analysis is not critical
Intervention	Strictly defined or standardised	Strictly defined or standardised
Implementation	optimised/tightly controlled provided by "experts" or "enthusiasts"	may vary—at discretion of clinicians provided by normal staff
Acceptance	optimised	Acceptance: variable
Control	Usually a placebo	Usually routine practice
Outcomes	Usually intermediate, clinical or laboratory or biological measures	Patient-centred, may include broader health-related quality of life measures
Analysis	Measures effect of actually receiving treatment as per protocol Analyse participants according to treatment actually received; i.e. for those who comply/complete	Intention-to-treat, according to the treatment allocated to rather than to treatment received Measures effect of making treatment available rather than actually receiving them
Settings	Well resourced	Variable resource levels

2.5.4. Measuring efficacy

a) Definition

The first level of research which this thesis undertakes consists of an assessment of the efficacy of the intervention. Efficacy refers to the effects of an intervention when delivered under ideal or experimental conditions. Efficacy studies are concerned with assessing if an intervention can work—i.e. whether it can do more good than harm—under ideal conditions [40, 97]. Further, efficacy trials aim to investigate how and how an intervention works. For this reason, they are also referred to as explanatory trials [40, 97]. Assessment of efficacy is considered a necessary step in the development and diffusion of a new technology or intervention [40]. If an intervention is shown to have no or negative effect under ideal conditions then it is unlikely to be effective in routine practice [40, 50]. On the other hand, once an intervention is shown to be efficacious, it is then useful to carry out an effectiveness trial to investigate if it can work in real-life settings; and implementation research to assess if all components can be delivered to acceptable standards in routine practice in a manner that is acceptable to the target audience [40].

b) Design of efficacy trials

Efficacy studies are designed to optimise performance or effects of an intervention. Therefore, an efficacy trial is characterised by (a) a well defined or standardised intervention that (b) is made available in a uniform fashion (c) within standardised and well-resourced settings (d) to specific target audience which (d) completely accepts, participates in, complies with or adheres to the intervention delivered [40, 98]. In addition, efficacy trials usually apply strict exclusion criteria so as to recruit participants with similar characteristics. Further, efficacy trials closely monitor the frequency with which interventions are applied and carefully measure outcomes of participants at various points in time [40, 98]. A typical study design for assessing efficacy consists of a randomised controlled trial (RCT) which involves (a) random assignment of participants to comparison arms, (b) conceal of allocation of study participants and (c) blinding of study participants and study teams to the interventions provided as well as blinding of outcomes assessment. Typically the comparison intervention consists of a placebo. However, for most public health interventions, the comparison intervention often consists of the best known or the current intervention. The assessment is then concerned with whether or not the new intervention does more good than harm when

compared with the intervention currently or previously in use. For practical reasons, test of efficacy is sometimes carried out using non- or quasi-experimental clinical trials which may not involve randomisation or blinding and may use historical controls, although they result in weaker causal inferences [40, 98].

c) Strength of RCTs

RCTs are considered the best design for attributing outcomes to an intervention. If well conducted, randomisation ensures that, on average, all potential confounders are equally distributed between comparison arms. Thus any significant difference between the study arms in the outcome assessed can be attributed to the intervention and not to a systematic difference between the two groups [50, 98-100].

d) Major limitations to RCTs

Threats to external validity

A typical RCT takes place under atypical (ideal) conditions characterised by standardised, well resourced settings, motivated research staff and participants, homogeneous population, and intense application and monitoring of the intervention [40, 98]. As such evidence from RCTs may not be generalisable to the general population in routine practice [98]. However, threats to external validity would be an issue if adoption decisions were based on results of efficacy trials only. According to diffusion of innovation model described above, adoption decisions are meant to be informed by evidence from effectiveness and demonstration trials. Efficacy trials are necessary but not sufficient for adoption decision-making [40]. Knowledge of efficacy is necessary to inform the decision as to whether or not it is necessary to carry out an effectiveness study and to aid interpretation of evidence from effectiveness trials. Therefore, in the context of this study, this limitation is irrelevant. Further, generalisable conclusions may be gained from systematic reviews and meta-analysis that identify similar effects in various populations [98]

Sample size, design effect and cost

The sample size required to detect a difference in effect between two groups is inversely proportional to the treatment effect squared [101]. Therefore, in order to detect small differences in effects, randomised trials require large sample sizes.

Many trials of public health interventions use cluster Randomised Controlled Trials (cRCTs) because individual randomisation is not feasible [98, 99]. If analysis is undertaken at individual level, sample size calculations for cRCTs need to take into

account the correlations among the individuals within the clusters (design effect) [98, 102, 103]. Randomisation can be effective in addressing baseline imbalance in cRCT if the number of clusters randomised in large. With fewer clusters, as is the case in most cRCTs, randomisation has less statistical power. The general rule of thumb is that more clusters with fewer individuals per cluster help to minimise the design effect [98]. Although use of stratified or pair-matched methods may minimise design effect in cRCTs, these methods also require large numbers of clusters that are similar enough to be paired or grouped [98].

Because RCTs require large samples, they are very costly to conduct—irrespective of whether they are individual or cluster randomised.

In assessing HW performance and guideline implementation, a common approach is to observe HW making observation at a few clusters (health facilities) [62, 90-92, 104-118]. Data on HW performance tend to be correlated because of the similarity in case management as all patients are often seen by only one or two HWs [102]. Randomisation of patients does not address the correlation in data on HW performance. Therefore, sample size calculation in such surveys need to account for potential design effect in the data on HW performance [102].

Contamination

Allocation and blinding is usually not practical or necessary in many public health interventions. Therefore, control conditions cannot be guaranteed: aspects of the intervention may be implemented in control settings. For example, in one cRCT designed to compare effects of implementation of an RDT-based guideline versus a conventional symptom-based guideline, some control facilities were later noted to have received and used RDTs provided through other supply chains [35]. Such contaminations can reduce the statistical power of the trial to detect effects of the intervention [98].

Ethics

Ethical challenges may arise by withholding potentially beneficial intervention from control population units, particularly if the intervention is associated with large positive effects [98].

Complex interventions

Many public health interventions are complex, consisting of several components [73, 74]. When interventions are associated with significant effects, it is important to identify which components of the intervention contributed to the outcomes. RCTs are unsuitable for attributing effects to components of interventions: it is difficult to identify the components of the intervention which is responsible for an observed effect [74]. If linked to outcomes assessment, implementation evaluation (demonstration studies) can be more useful in explaining positive, modest and insignificant results [73, 74].

e) Choice of a systematic review of RCTs to assess efficacy

A systematic review is a method of systematically collating all empirical evidence that fits pre-specified eligibility criteria in order to answer a specific research question. It uses explicit inclusion criteria that are selected to minimise bias and increase internal validity of the findings [119].

A systematic review was chosen because it was not feasible to undertake an RCT in the context of this thesis due to financial and time constraints. Secondly, a systematic review pools together all possible evidence from a broader context. Therefore, it represents the best evidence for decision making at a broader level. Although RCTs are criticised for its limited external validity, systematic reviews can lead to generalisable conclusions by identifying similar effects in various populations [98].

Evidence from non-randomised trials was excluded from the review. Selection biases (confounding) are likely to be greater for non-randomised studies than for RCTs [119]. Inclusion of evidence from non-randomised trials in a meta-analysis can lead to a shift in the estimate of the effect of an intervention (systematic bias) and excessive heterogeneity among studies [119].

Potential limitations

The research quality criteria used to select studies for inclusion in a systematic review of effects of interventions has been criticised for being overly biased against studies that may show large effects of an intervention, but whose designs are deemed to be sub-optimal as per the selection criteria [73]. On the hand, it may be biased towards interventions with marginal effects because the designs of the studies used in evaluating effects meet the selection criteria.

Many studies used in evaluating public health interventions employ cluster randomised trials in which analysis is done at the level of the individual patients, leading to unit of analysis error which may lead to over-precise results. If included in a meta-analysis without correcting for clustering, estimates of effects from cluster randomised trials may carry more weight on the pooled result of the meta-analysis and lead to a biased estimate [73].

2.5.5. Measuring effectiveness

a) Definition

The second level of study applied in this thesis is an effectiveness trial. Effectiveness refers to the effect of an intervention under routine or "real-life" conditions. Effectiveness trials aim to demonstrate whether an intervention does more good than harm in routine or "real life" circumstances. For this reason, effectiveness trials are also called "pragmatic" trials because the focus is assessing if the intervention does work in the real world [40, 97].

b) Design of an effectiveness (pragmatic) trial

A typical design of an effectiveness trial would incorporate situations the clinicians are likely to encounter in usual practice [40, 97]. For example participants' selection would be broad in order to assess if the intervention works for the general population. The intervention would be provided by normal program staff, who may not receive any special incentives to deliver the intervention as defined. Delivery of the intervention is normally at the discretion of the program staff and may vary within and between sites. Therefore the magnitude of effects of the intervention may vary between sites due to variability in delivery and quality of the intervention to the target audience, and/or variability in acceptance of (participation in, compliance with, or adherence to) the intervention by the target audience [40]. Control interventions usually consist of current or previous interventions or treatment.

It follows that a good effectiveness study would include an assessment of implementation (availability, quality, and acceptance) as well as measuring of effects of the intervention. Without assessment of implementation, and especially without prior knowledge of the level of efficacy, it is difficult to determine whether a lack of intervention effects is due to inefficacious intervention or inadequate programme implementation [40].

Randomisation, allocation concealment and double blinding are usually much more difficult to implement in less-controlled, real-world settings [40, 98]. Quasi- or non-experimental designs are used more often in assessing effectiveness of public health interventions than the traditional RCTs described above [120, 121]. Examples of quasi-experimental designs recommended for effectiveness trials include:

(a) *basic pretest-posttest design*: this involves comparing estimates of effects of the intervention in several populations before the intervention, with estimates of effects in the same population after they have received the intervention. Maturation and history are major threats to internal validity of this design [122, 123]. Maturation refers to internal changes in the population over time while history refers to exposure of the population to external events (external to the intervention) that may affect their post-test estimates of effects [122, 123].

(b) *Control-group pretest-posttest design*: this design consists of an intervention arm which essentially involves pre-post comparison as described in (a) above, and a parallel control arm, which also involves a pre-post comparison at the same time points but without application of the intervention. Introducing a control arm is useful in assessing the role of maturation and history. Populations may be assigned to the two arms through random allocation (and unit of allocation are usually clusters) [98, 122, 123]. Additionally the comparison groups may be matched or unmatched [98, 122, 123]. Control-group pretest-posttest is the design used in the trial analysed in chapter 4 of this thesis.

(c) *Intervention versus control group comparison*, without a baseline survey, which may be randomised, or non-radomised, matched or un-matched. Baseline imbalance and confounding would be the major threats to this design in the absence of randomisation and/or matching.

2.5.6. Measuring implementation

a) Definition

The third level of research proposed in this thesis consists of a demonstration or implementation study. A demonstration study consists of routine implementation of an intervention of proven efficacy and effectiveness in whole communities (e.g. a district, region or state), and monitoring the implementation [40]. The focus of a demonstration study is on implementation of the intervention (what components, what quality), acceptability (participation or adherence) and context (by whom, under what conditions) [73]. However, demonstration studies may include an assessment of outcomes in the population (morbidity, mortality, behaviour change, cost, etc) [40].

During demonstration, delivery of interventions may vary between settings, or can be varied deliberately between sites in order to determine which model is more effective in delivering the intervention, or has a better reach. Therefore, demonstration studies can show the extent to which various elements of a multi-component intervention are actually implemented [73, 74]. In addition, it enables the identification of facilitating and inhibiting contextual factor [73]. If linked to outcomes assessment, demonstration studies can be useful in interpreting effects of a complex or multi-component intervention. Collection of process and context variables helps to explain how and why an observed change has occurred [40]. Although it may be difficult to attribute effects directly to single and specific element of a multi-component intervention [40, 73], detailed implementation evaluation can be used to generate hypothesis about causal relationship between intervention components and specific outcomes [40, 73, 74].

b) Design and methods

Studies evaluating the quality of clinical practice or guideline implementation usually employ one or a combination of the following methods (a) observing a HW perform a set of clinical tasks which are specified in the guideline [5-23], (b) "gold standard examination" of study participants by a more qualified research staff [8, 13, 16, 18, 24], (c) interviewing the HW (e.g. about training, knowledge, attitude, experiences) [6, 8, 10, 16, 18, 25-27], (d) exit interview of participants (e.g. to evaluate clinical notes, diagnosis and prescriptions made) [6, 8, 10, 13, 16-18, 27], and (e) assessing the adequacy of essential health systems inputs and support services that are required for effective implementation of the guideline [6, 8, 10, 26].

For guidelines originating from international organizations and networks—such as the WHO, UNICEF, INRUD—studies typically use standardized and field-tested indicators derived from these organisations [28-30]. These indicators are normally adopted wholesome or are adapted slightly to suite the context of the studies. In a few instances, indicators have been developed by the researchers themselves, especially where guidelines were developed locally [13, 17]. Studies normally vary in the scope of items assessed or how indicators are measured, even when common tools have been used.

c) Measures of HW performance

Measures of HW adherence to guideline have commonly taken the form of point estimates (with associated estimates of errors) [6, 7, 9, 15, 16, 18, 24, 26] or categorical (qualitative) descriptions [20-23], or both [8, 10, 13]. Where both point estimates and categorical outcomes are provided, the latter tends to refer to the quality of prescribing (type of medications and regimen). When indicators are measured in the form of point estimates, they are frequently expressed as adherence or performance scores, which represent the percent of expected tasks (i.e. those specified in the guidelines) which are performed by the HWs [8, 10, 13, 15-18, 26]. Usually, this process involves counting the number of expected tasks performed by the HW, without taking into account the quality of execution. Occasionally, or additionally, point estimate measures have been expressed in the form of average values (±standard deviation), representing the number of expected tasks performed by the HW per consultation. This form of measure has commonly been used in evaluation of guidelines focusing on rational drug use [7, 31-34]. Occasionally, they have also been used in assessing IMCI-related guidelines [19].

The major limitation of using point estimates as measures of HW adherence to guidelines is that frequently standards for judging adequacy of HW performance are not provided. In the absence of a standard, it is difficult to judge if a given level of performance (percentage point or average value) represents acceptable quality or poor quality. Notable exceptions are studies assessing implementation of guidelines on rational drug use, for which standards are normally provided. These standards are based on surveys carried out in developing countries in the 1990's (INRUD). Therefore, most of these standards have become obsolete because many clinical guidelines in developing countries have evolved since the 1990s, especially those in malaria and HIV endemic countries [9, 19, 124].

Much more recently, LQAS-based methods have been widely used in assessing HW performance descriptions [20-23]. LQAS-based methods often apply the same data collection techniques and indicators described in (b) above. However, the LQAS-method is a triage system which allows judgement to be made about the adequacy of HW performance against pre-set performance standards. The outcomes in an LQAS-based survey are dichotomous; e.g. either acceptable or unacceptable, low or high, adequate or inadequate, etc. The LQAS-method allows the identification of poorly performing HWs or areas that require urgent action. Therefore, the LQAS method uses an explicit approach in assessing HW performance, and is action oriented. LQAS principle and its application in healthcare are described further below.

2.6 Use of LQAS method to assess quality of care

Chapter 5 of this thesis uses Lot Quality Assurance Sampling (LQAS) method to assess whether districts are adequately embracing the RDT guidelines in routine clinical care of febrile children.

2.6.1 Basic principles

a) Description

LQAS is a quality control method that originated in industry in the 1920's, adapted to the health sector in the 1980's. LQAS was developed to classify the quality of a particular batch, or lot, of goods as acceptable, or not, according to a specified performance standard. This is done by counting the number of 'defects' in a small sample of a batch. The entire lot, or batch, is then rejected or accepted depending on the number of defects in the sample. The maximum number of defects permitted in a sample considered to be of good quality is referred as the decision rule (d). The outcome in LQAS-based assessments is binary; e.g. either "acceptable" or "unacceptable". Hence the LQAS is a classification method and is typically not used to calculate point estimates.

LQAS appears sensitive to detecting poorly performing parts of the system being sampled. The aim is to help local managers to identify components of the system that require urgent action and to prioritise resources [104, 125].

The main advantages of LQAS method are that it requires a small sample size, is rapid and therefore it is not resource-intensive [105-107, 126]. The trade off in using small sample sizes is that some lots may be rejected even if they are in reality acceptable, while others with higher than acceptable levels of defects may escape detection [125]. Further, because LQAS samples are small, the 95% confidence intervals for point estimates are wide. However, by aggregating data across different lots, a point estimate for a given indicator can be calculated due to the large total sample size [107].

The LQAS-based methods have recently been applied extensively in developing countries to assess child survival and maternal and child health interventions [104, 106, 127-129], to detect malaria epidemics [130], to assess communities for *Schistosoma* [126, 131] and to identify strategies for control of yaws[132]. It is potentially useful for malaria programme management [125, 133].

b) Standards

To classify a performance as either acceptable or unacceptable, four parameters are set beforehand. The first is the level of the desired performance which defines acceptable quality, also referred to as a performance threshold. The second is the level of performance that is considered to be seriously below the performance threshold and is deemed "unacceptable", at which point managerial attention is a priority. The third is the level of risk one is willing to take in judging a lot as having achieved the performance benchmark when in fact it has not (β-error). The fourth is the level of risk one is willing to take for judging a lot as having failed to achieve the performance benchmark when in fact it has (α-error) [104-107]. The typical decision rule as used in public health programme assessments is one in which both α and β <0.10 and their sum is <0.20[104].

These benchmarks are usually set in consultation with programme managers [104]. For example, in a previous LQAS-based survey of immunisation coverage, the managers chose a target of 80% to define adequate immunisation coverage. This target is called the upper threshold. Further, they chose coverage level of 50% to denote a highly "unacceptable" coverage level which the LQAS was designed to detected with low error [106].

c) Sample size and decision rules

LQAS is a statistical method that uses cumulative probabilities to calculate an appropriate sample size and a decision rule to reliably identify production units (e.g. health facilities in a district) that are performing at a desired level of quality [104]. The standard approach is to use the binomial formula when dealing with a large population size, e.g. when calculating the number of observations (patients) required to judge the adequacy of clinician adherence to guidelines [104, 106]. The formula for the binomial distribution is shown in Appendix 1. When dealing with a small and finite population, a hyper-geometric formula is applied [104]. For example, in calculating the number of facilities to sample in a LQAS-based study in a given district, a hyper-geometric formula is more appropriate since the total number of facilities (universe) is known and is small. The binomial and hypergeometic probabilities approximate each other in large populations but differ in small ones.

The binomial or hyper-geometric formula calculates the probability of selecting a certain number of defects "d" (e.g. unvaccinated children) from a sample "n" drawn from a lot of a certain quality "P" (e.g. health area with a given immunisation coverage). Conversely, the binomial or hyper-geometric formula calculates the probability of correctly identifying a lot of quality "P" given a level of "d" defects in a sample "n" drawn from the lot [104].

The following example from Valadez (1991) illustrates how LQAS and decision rules work [104]. In assessing adequacy of immunisation coverage among children in a given community, assume 80% and 50% are chosen to denote adequate coverage and unacceptable coverage respectively. Assume the calculation yields a sample of 19 children and a decision rule of 6 (i.e. 19:6). The decision rule 19:6 means that when 6 or fewer unvaccinated children are detected in a sample of 19 children, the sample, and the community from which it is drawn, is classified as having achieved adequate immunisation coverage. This decision is associated with an error (β-error), which represents the risk of classifying the coverage in the community as adequate when in fact it is not (i.e. when the actual coverage is less than 80%).

The binomial and hyper-geometric formulas calculate the exact risks associated with applying the decision rules. When applied to the binomial formula, the 19:6 rule shows that a community with 80% coverage will be classified as adequate 93% of the time [104]. That is, 93% of the time the number of unvaccinated children (defects) in a sample of 19 children will be 6 or fewer; 7% of the time, the number of unvaccinated children in the sample will be more than 6, and the sample—and the community from which it is drawn —will be misclassified as having achieved inadequate coverage. If the same rule is applied to a community with 50% coverage, it will be incorrectly identified as having achieved adequate coverage—that is, the samples will contain 6 or fewer unvaccinated children—8% of the time [104]. In 92% of cases, a sample drawn from a population with 50% coverage will contain more than 6 unvaccinated children and will be classified as having been inadequately covered.

If the same 19:6 rule is applied to a sample from a health area with coverage between 50% and 80%, it will be classified as adequate or unacceptable depending on how close the coverage falls to the upper and lower bounds of the triage system. The classification errors (α and β errors) associated with the middle grounds are higher than the

corresponding values at the upper and lower thresholds [104]. Therefore, LQAS optimises the identification of extreme performance: worst of worst, which require priority attention and/or best of best, so that resources are not wasted.

2.6.2 LQAS for Health Facility Assessment

The LQAS method has been applied in the health sector to determine whether a particular health programme, or health facility, meets a desired performance threshold [104-106]. In Health Facility Assessment (HFA), LQAS is used in a 2-stage sampling design. Firstly, LQAS is used to determine whether an acceptable proportion of health facilities in a supervision area perform a given clinical task adequately. Typically, at least 80% of health facilities are expected to have clinicians displaying behaviour consistent with clinical guidelines (upper threshold) [104, 106]. Additionally, LQAS is set to detect supervision areas with higher error rates in which, typically, only 50% or fewer of the health facilities exhibit the appropriate behaviour (lower threshold) [104, 106]. Secondly, LQAS is used to determine the number of clients needed to judge adequacy of a single provider's performance of a given clinical task. Previous studies have typically used an upper threshold of 95% to define adequate performance—expecting a given provider to deliver specified services using the correct technique at least 95% of the time. A performance threshold of 50% has typically been used to denote unacceptable health worker performance [104, 106, 127, 128].

2.7 Overview of malaria diagnostic methods

2.7.1 Clinical diagnosis

Prompt and accurate diagnosis of malaria is the mainstay of the WHO malaria control strategy. Despite considerable progress in malaria control over the recent years, diagnosis of malaria remains problematic due to lack of diagnostic techniques that are technically, operationally and financially sustainable at primary care settings in endemic countries [134]. Consequently, in most endemic countries, the diagnosis of malaria is based on clinical judgment in most health facilities [15, 22].

Clinical features, including fever, are poor predictors of malaria disease. Several studies from different epidemiological settings show that 32% to 93% of patients classified as malaria on clinical grounds alone are slide negative (mean: 61%) [11-14]. On the other hand some patients with malaria parasites, or non-malarial febrile illnesses (NMFIs), such as acute respiratory infection, could be left untreated or are maltreated. This increases the potential for resistance to the medicines used, or may result in complicated forms of the conditions (malaria or NMFIs) due to improper or lack of treatment. The patient and the healthcare system may incur unnecessary and substantial costs as a result of such misdiagnosis [11-14].

2.7.2 Biologic methods

Research in malaria diagnostics has been limited. Microscopic examination of Giemsa-stained blood smears was introduced more than 100 years ago and has been used as the gold standard technique for malaria since [20]. In the past 50 years, alternative parasite-based techniques have emerged. They include enzymatic immunoassays (ELISA) and immunofluorescence antibody assay (IFA) [20, 22, 134]. These tests detect antibodies to the asexual forms of malaria parasites which appear days to weeks after the infection and may persist for months. Therefore, they can not distinguish between current or past infections, and are therefore of limited value in guiding treatment of malaria [134].

Antigen-based methods have also become available. The most significant of these is the immunochromatographic assay, which forms the basis of the current commercially available malaria RDTs [20, 134].

Much more recently, polymerase chain reaction (PCR), a molecular method that uses DNA probes, has also been introduced. PCR techniques are much more sensitive than, or as sensitive as, microscopy [20, 22, 134]. However, PCR is costly, difficult to run and requires sophisticated equipment and highly qualified technicians. Therefore, currently the method is used mainly for research [134].

Other methods which have emerged recently include fluorescent staining techniques such as the quantitative buffy coat (QBC) analysis and acridine-orange staining of thin blood smears [20, 22, 134]. Although these techniques appear to have high sensitivity at low parasite density (100 parasites/ μL) [135], they are, so far, of limited value in clinical settings for the same reasons cited above [22]. Use of deplorased laser light to detect malaria pigments, and mass spectrometry have shown limited success [20]

2.7.3 Microscopic examination

Microscopic examination remains the gold standard diagnosis of malaria. Microscopy has several advantages. It is inexpensive to perform, able to differentiate malaria species and quantify parasites [20]. Expert microscopy of Giemsa-stained thick blood film is sensitive at low parasite density (5 to 20 parasites/μL) [20-22]. It is useful in parasite quantification, species differentiation and staging of parasites, all of which are useful indicators in selection and evaluation of treatment. It can also be used for diagnosing other conditions. Furthermore, malaria smears provide a permanent record that can be used in quality assurance of malaria diagnosis [22].

However, microscopy has several limitations. It is labour-intensive and time consuming. Routine clinical microscopy can not detect very low parasitaemias (<5 to 10 parasites/μL)[22]; the detection threshold is estimated at 50 to 100 parasites/μl, much higher in settings with less skilled microscopist [20, 22]. Mixed infections are often missed (requiring thin films).

In addition, widespread deployment of microscopes in rural health centres, the first contact points for most febrile patients, is not feasible in the near future because of the high cost of procuring sufficient number of microscopes, the operational costs, and the technical capacity building required [23, 24].

Because of the operational and technical limitations of microscopy, Rapid Diagnostic Test (RDT) is now recommended as a quick and accurate parasite-based technique for

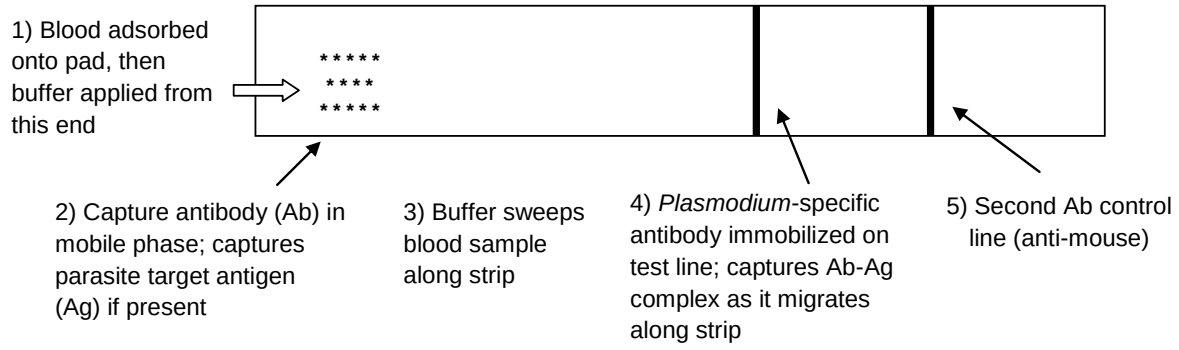
use where high quality microscopy is not available [14, 19, 53]. Malaria RDTs have been widely tested over the last 20 years [20, 22]; and RDT-based guidelines are now operational in many malaria endemic countries to support the diagnosis and treatment of febrile patients [25].

2.8 Malaria rapid diagnostic tests

2.8.1 Technology

An RDT is a device that is capable of detecting malaria antigens in a small volume of blood, usually 5 to 15 μL. It is a method based on immune-chromatographic assays. The modern malaria RDT consists of antibodies against malarial antigens impregnated on a nitro-cellulose strip encased in a cassette [20, 22]. The clinical sample migrates along the strip by capillary action (Figure 5 below).

Figure 5: Schematic of how modern malaria RDT works



Source: Dorsey and Hopkins [81]

Two types of antibodies are used for any target antigen: a capture and a detection antibody. The capture antibodies serve to extract and bind the parasite antigens from a migrating liquid sample. The detector antibody, joined to an indicator, combines with the immobilised target antigen to produce a bright coloured light, usually appearing in 5 to 20 minutes [20, 22].

Three types of antigens are commonly targeted by the malaria RDTs which are currently available.

Histidine-rich protein-2 (HRP-2): this is a protein unique to *P. falciparum*. It is produced predominantly by the asexual form of the parasite. It is also found in gametocytes which can be detected in blood in the absence of the asexual form that causes disease [22]. HRP-2 can be detected at parasite concentration that is lower than is possible with other target antigens (i.e. <100 parasites/μL). HRP-2 antigen is known to persist in blood for

more than 28 days, well after the asexual forms of the parasite that cause disease have cleared from the blood. Therefore HRP2-based RDTs are not useful for monitoring response to treatment [20, 22].

Plasmodium lactate dehydogenase (pLDH), which is an enzyme produced by all species of *Plamodium*. It is produced by both asexual and sexual forms of the parasite. Antibodies have been developed that can detect pLDH that are unique to *P. falciparum* or *P. vivax;* or pLDH which is found in all *Plasmodium* species (pan malaria). The level of pLDH falls rapidly after initiation of therapy, clearing from blood within one week [20, 22]. Therefore pLDH-based RDTs are potentially useful for monitoring response to therapy [20].

Aldolase: this is also an enzyme, produced by all human species of *Plasmodium* parasites and is used as a pan malarial antigen target. It is produced by both asexual and sexual forms of the parasite. The level of aldolase also falls rapidly after initiation of therapy [20, 22]; hence, aldolase-based RDTs are also potentially useful for monitoring response to therapy [20].

No evaluations of *P. malariae*-specific or *P. ovale*-specific antigens have been published.

2.8.2 Commercial types

Commercial malaria RDTs carry 3 major combinations of antibodies to suite different malaria epidemiological contexts:

a) *HRP-2-based RDTs*: the vast majority of RDTs which are recommended for use in sub-Saharan Africa carry anti-HRP-2; hence they are specific to *P. falciparum* [20, 22].
b) *pLDH-based RDTs* carry antibodies targeting pLDH produced by either *P. falciparum* or *P. Vivax;* or *pLDH* produced by all *Plasmodium species (pan-malaria) [20, 22]*.
c) *Combination-type RDTs,* are essentially HRP-2-based; in addition, they carry other antibodies targeting *P. vivax*-specific pLDH, pan malaria pLDH or aldolase, thus making them capable of detecting mixed infections, or differentiating between species [20, 22].

2.8.3 Factors influencing performance of RDTs

Several factors in the manufacturing process and those in the environment may affect the diagnostic accuracy of malaria RDTs. Table 5 below describes the main factors which may influence the performance of malaria RDTs.

Table 5: Factors influencing performance of malaria RDTs

FACTOR	EXPLANATION
Test devices	
Manufacturing	
Lot used	The same product from the same manufacturer may vary in test properties because of material modifications over time
Transport and storage conditions	
Humidity	-Humidity rapidly degrades RDT
Temperature	-Higher temperatures degrade RDTs by altering the flow characteristics of the nitrocellulose wick, deconjugation of the capture antibody-dye complex, detachment of capture antibody from wick -Drastic temperature changes (freeze-thawing), e.g. during transportation, can have similar effect
Preparation and interpretation	
Blood buffer	-Controls flows, sometimes lysis
Blood volume	-Inadequate volume reduces available antigen. Excess volume inhibits clearance of blood stain, reducing clarity of test lines
Storage & duration	-Lysis of red blood cells occurs during storage, which results in decrement of antigen activity over time
Visual acuity	-Test line can be faint at low antigen activity (long storage duration, low parasite density); it might be difficult to see in poor light, or if the reader has poor visual acuity
Patient and parasite	-Sequestration of parasites determines antigen concentration and parasite density in peripheral blood -Antigen production varies with parasite life cycle and between parasites -Antigen structure vary between and with parasite species and strains -Parasites recovered directly from patients may show different antigen activity from that recovered from laboratory cultures -Lysis of red blod cells and aggregation of parasitized red blood cells can reduce consistency of flow
Training	-the quality of RDT prepared by well trained laboratory workers (evaluation studies) might vary from that prepared by field workers with limited training
Reference standards	
Microscopy or PCR	-Poor sensitivity of reference standard reduces apparent RDT specificity -Poor specificity of reference standard reduces apparent RDT sensitivity
Study population	
Parasite factors	-Parasite density affects sensitivity -Parasite prevalence affects predictive values
	-Antigen activity varies between wild and cultured parasites (laboratory trials)
Patient factors	-Treatment history and effectiveness of treatment varies between patients -Presence of substance prone to cause false-positive reactions (e.g. Rheumatoid factors) can vary between patients

For example, commercially distributed RDTs commonly undergo material modification after their initial introduction [22, 136]. Thus different batches of the same product from the same manufacturer may behave differently in both operational and performance characteristics [22, 136]. Further, because RDTs are based on antigen-antibody interactions, they deteriorate very rapidly on exposure to high temperature and humidity. Manufacturers commonly specify storage between 2°C and 30°C. In practice RDTs are frequently exposed to >30°c and/or 70% humidity in clinical settings in the tropics [20, 22, 65, 136]. Product evaluation trials often use well trained well-trained laboratory personnel. In clinical settings, end-users are typically health care workers with limited training [20, 22, 65, 136].

2.8.4 Performance of RDTs in field trials

The diagnostic accuracy of the common commercially available RDTs has been extensively investigated under different transmission settings. Overall, the results show considerable inconsistency in the diagnostic accuracy of RDTs [20, 22, 135].Table 6 below shows the results from an extensive systematic review of RDT performance from trials conducted in clinical settings in endemic countries [135]. It shows that HRP-2-based RDTs are generally more sensitive than pLDH-based RDTs when used in febrile patients [135]. Additionally, it reveals that the average sensitivity of both types of RDTs is below the threshold of 95%. On the other hand, it shows that the specificity of both types of RDTs is generally higher than the recommended threshold of 95% [135].

Table 6: Sensitivity and specificity of two RDTs in clinical settings in endemic countries

		Diagnostic accuracy relative to microscopy			
RDT type	Number of trials reviewed	Sensitivity (95% CI)		Specificity (95% CI)	
HRP-2	13	92.7	(91.0-94.5)	99.2	(98.2-99.9)
pLDH	6	67.1	(62.8-71.3)	98.4	(97.5-99.6)

Source: Ochola 2006 [135]

Reviews by Wongs in 2007 and Murray in 2008 also show that the sensitivity and specificity of HRP2-based RDTs vary widely in clinical settings where the test is performed on febrile patients by trained health workers. Sensitivity was found to vary between 85 to 100%; and specificity varied from 90% to 100% [20, 22].

Inconsistency in results of field trials are usually attributed to design limitations and/or may be due to a combination of the factors outlined above [22, 65, 137]. These inconsistencies hamper comparability of trials results and broader applicability in routine settings [22, 65, 137].

2.8.5 WHO product testing

Inconsistency in results of field trials has led to the introduction of a WHO-led collaborative initiative of RDT product testing and quality assurance [20, 22, 65]. This initiative aims to standardise the assessment of RDT performance, to guide procurement decisions, and to drive improvement in the quality of manufacturing [20, 22, 65]. Therefore, guidelines and standards have been introduced in order to standardise field components of sensitivity/specificity trials of malaria RDTs [65, 68, 69]. Additionally, limited laboratory-based performance analyses of commercially available RDTs are conducted by testing assays against panels of malaria parasite-infected blood samples. The evaluation is designed to provide comparative data on the performance of submitted production lots of RDTs; and to give an indication of which products are likely to give higher sensitivity in the field, especially in populations with low-density infections [22, 65].

Product testing involves assessing each product in terms of (a) positivity rate (panel detection score), (b) false-positivity rate, and (c) heat (thermal) stability, and (d) ease of use [65, 69]. So far three rounds of product testing have been completed since 2006. Over the 3 rounds, several RDTs have demonstrated consistently high positivity rates at low parasite densities (200 parasites/ μL) [65]. Further, they have demonstrated low false positive rates, stability at tropical temperatures, easy-of-use, and ability to detect *P. falciparum, P. vivax* infections or both [65].

Laboratory-based studies do not fully depict the complexities and physical stresses to which a diagnostic test may be subjected to under field conditions [22]. Therefore, standardised field trials are still necessary. Further test performances vary between lots,

and widely between similar products, despite standardisation of procedures [65]. Therefore, the WHO recommends lot-testing post-purchase and prior to use in the field [65].

Because of the likelihood of inter-lot variability, the WHO strongly recommends that samples from each production lot should be tested pruior to dissemination to the field. Accordingly, additional to product testing initiative, the WHO, TDR and FIND support national programmes in assessing RDT lots prior to purchase [65].

.

2.8.6 Factors influencing utility of RDT-based policies in practice

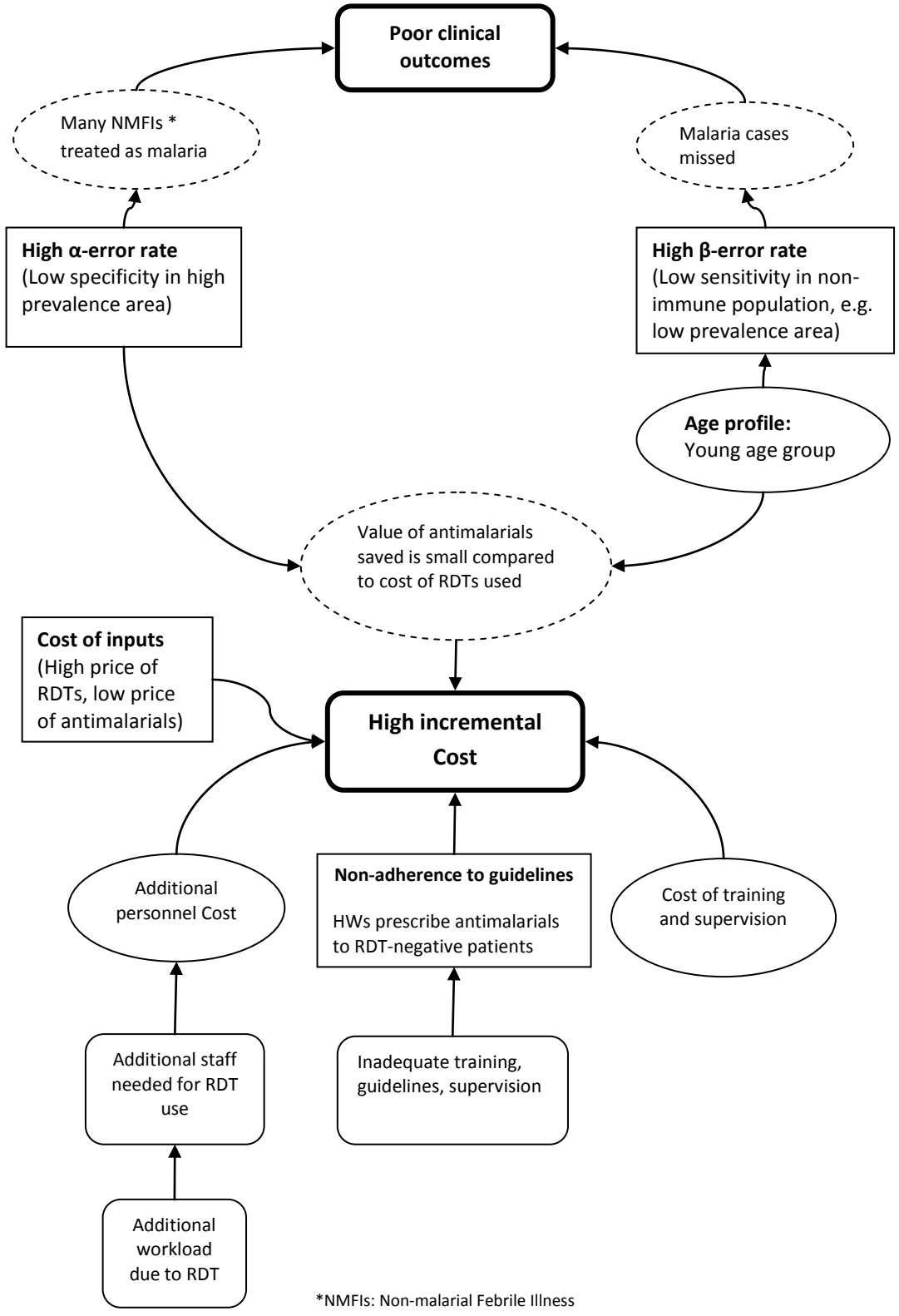
Several factors interact to determine the incremental cost and clinical outcomes of RDT-based policies relative to presumptive treatment of fever. The main factors include the diagnostic accuracy of RDTs in different malaria transmission settings, age profile of the target population, health worker adherence to RDT-based guidelines and prices of inputs—especially prices of RDTs relative to those of antimalarials used in treating uncomplicated malaria [21, 22, 138, 139]. Figure 6 displays how the various factors might interact to influence clinical outcomes and incremental cost in settings where RDT-based policies have replaced or might replace presumptive treatment of fever. There is need to quantify the benefits and incremental cost of RDT-based policies using analyses which take into account all these factors, and which use actual field data. Such analyses could provide more accurate insight into the value of RDT-based policies in different settings.

a) Diagnostic accuracy of RDTs

Sensitivity

To be useful diagnostic tools, malaria RDTs in sub-Saharan Africa must achieve sensitivity for *P. falciparum* of at least 95% at parasite density of 100 parasites/μL of blood, or higher [20, 22, 139]. Clinical disease may occur at lower parasitaemia in a non-immune population, such as those in low transmission settings and/or children under 5 years of age [22]. Therefore, RDTs used in non-immune populations are required to maintain high sensitivity (at least 95%) at a concentration of 50 parasites/μL of blood [22]. The sensitivity of the currently available RDTs tends to fall far below 95% at parasite density lower than 100 parasites/μL of blood. Therefore, there is increased risk of morbidity or mortality from missed malaria cases in non-immune population. This is a key point of contention against scaling-up the use of RDT-supported guidelines in areas where presumptive treatment is currently the norm, especially for use in children, until sufficient evidence is provided to support their effectiveness [26, 138].

Figure 6: Determinants of effectiveness and incremental cost of RDT-based policies



Specificity

It is recommended that malaria RDTs must achieve specificity of at least 95% at parasite density of 100 parasites/μL of blood, or higher [20, 22, 139]. In high transmission settings, RDTs have lower specificity, particularly the HRP-2 type which is recommended for endemic high transmission countries [21, 22, 138, 139]. In these settings, febrile illnesses are commonly accompanied by parasitaemia and antigenaemia that are not related to the illness; and which may persist after initiation of anti-malarial therapy [21, 22, 138, 139]. Therefore, the number of false positive diagnoses averted in high transmission settings by a highly specific test is marginal. Accordingly, the cost of RDT per case of fever could exceed the drugs cost saving per case [21, 22]. Therefore, although a standard specificity of 95% is generally recommended [21, 22], lower specificity (>90%) is generally acceptable in high transmission regions [139].

Because RDTs have low specificity in high prevalence areas, many Non-malarial Febrile Illnesses (NMFIs) are treated as malaria. The actual agents causing the fever may be left untreated, and the patients may deteriorate, because they are malaria parasite positive [138]. Therefore, use of RDTs in high transmission settings may not be beneficial from both economic and clinical stand points.

In low prevalence settings, use of RDTs can avert a significant number of false positive diagnoses and overtreatment with ACT, which characterise presumptive treatment. Therefore, high specificity (at least 95%) is a necessity in low transmission areas [21, 22].

b) Age group

In children under 5 years of age, clinical disease may occur at parasite densities which are much lower than what the available RDTs are capable of detecting (e.g. < 50 parasites/μL of blood) [22]. Therefore, compared to symptoms-based treatment, use of RDTs to guide treatment in febrile children could result in increased morbidity or mortality from missed malaria cases.

Further, the amount of antimalarials saved in febrile children by the use of RDTs is much smaller than the amount saved in older patients. Therefore, the incremental cost of using RDTs in children may exceed the average drug cost savings, particularly in high prevalence areas [12, 80, 140]. Therefore, the net incremental cost of introducing RDTs in a particular setting depends on the age profile of the population in that setting. If the users of care in the setting are predominantly younger, RDT-based policies may add

cost to the health system relative to presumptive treatment [12, 140]. Analysis of incremental cost and effectiveness by age-group is useful in deciding whether or how to target RDT-based policies.

c) Prices of inputs

The net incremental cost of introducing RDT-based policies depends on the price of RDTs relative to the price of the antimalarials used for treating uncomplicated malaria. The higher the cost of RDTs relative to the price of antimalarials used, the less likely will the use of RDTs in treating febrile patients save any drug costs [12, 13, 141]. The recommended cost per test of RDT is US$ <1 for endemic countries [21, 22]. In 2010, the average price of P. *falciparum*-specific (HRP-2) RDTs was US$ 0.51 (range: 0.42 – 0.88) and US$ 0.69 (range: 0.58 – 1.05) for multi-species test [142]. Two main types of ACTs are currently used in treating uncomplicated malaria in endemic countries: artemether-lumefantrine (AL) and artesunate-amodiaquine (AS-AQ) [142]. In 2011, the average price of an adult dose of AL was US$ 1.30 to US$1.40; and the average price of an adult dose of AS-AQ was US$ 0.78 to US$ 0.94 [142].

2.8.7 Economic evaluation of RDT-based guidelines

This section reviews the economic evaluation studies that have thus far been undertaken to assess the economic potentials of applying RDT-based guidelines to manage febrile patients instead of presumptive treatment of cases.

a) Definitions

Full economic evaluation involves comparing the cost of interventions with their effects or outcomes. It takes the form of cost-effectiveness analysis (CEA), cost-minimisation analysis (CMA), cost-utility analysis (CUA), or cost-benefit analysis (CBA). In CEA, outcomes are described in their natural or physical units, e.g. number of deaths averted, number of patients cured. CMA is a form of CEA; it is undertaken when the outcomes of comparison interventions—e.g. the number of death averted—are the same or similar in magnitude. The outcomes in a CUA have two dimensions: length of life and the quality of life gained by an intervention. In CBA, outcomes are valued in financial terms [143].

Partial economic evaluation takes many forms [143],

a) *Cost analysis*, where interventions are compared in terms of cost only; costs are measured per units of outputs (e.g. number of participants diagnosed, treated etc) rather than per units of outcomes.
b) *Effectiveness analysis,* interventions are compared in terms of their health outcomes only
c) *Cost-effectiveness description,* where the cost of an intervention is compared with its outcome, in the absence of an alternative or comparison intervention
d) *Cost description,* showing only the cost of an intervention, in the absence of a comparator
e) *Effectiveness description,* showing only the health outcomes of an intervention, in the absence of a comparator

Therefore, partial evaluations do not give the full picture of the economic worth of an intervention or comparison interventions.

b) Papers and search strategies

The search strategy for this section is shown in Appendix 3 (c), search set 24. The review identified 4 CEA [72, 140, 144, 145] and 3 cost analysis papers[4][12, 13, 33], evaluating the effects of RDT-supported treatment versus presumptive treatment of fever. One CEA study was excluded from the review because the method was not clear [145]. Of the 6 studies included in the review, only one (cost analysis) used patient-level data from a randomised trial [33]. Two cost analysis studies used population-based (summary) data from on-going programmes [12, 13]. The remaining 3 (CEA) involved the use of models which used estimates of costs and outcomes from literature and expert opinion [72, 140, 144]. Table 7 summarises the main features and findings of the studies reviewed.

c) Main features

These microeconomic studies vary in design and range of inputs included in their analyses. All studies are based on outpatients presenting with uncomplicated fevers and are carried out in the context of ACTs. All analyses assume providers' perspective. All studies, except one, assume perfect clinician adherence to guidelines; the exception is the study by Zurovac et al in 2008, which incorporates actual adherence data from a randomised trial [33]. All studies analyse variation in costs and outcomes under different scenarios of malaria prevalence, participant age, types of ACT and (in some studies) and the diagnostic accuracy of the RDT used in the study.

d) Main findings

Cost-saving

All studies measured the degree of cost reductions attributable to the use of RDTs under different scenarios. The results suggest that use of RDT to guide the management of fever can reduce the cost of diagnosis and treatment by 21% to 25% in low to moderate malaria prevalence areas (<50% prevalence) [12, 13, 33, 72, 140, 144]. This is conditional on health workers adhering with the guidelines (using the test and treating according to the results) all or most of the time. The analyses also show that use of RDT-guideline is unlikely to be cost-saving in high prevalence areas even if health workers comply with the guideline [12, 13, 33, 72, 140, 144]. All the analyses show that

[4] All the cost analysis papers were reported as cost-effectiveness analyses

RDT-supported treatment is not cost-saving for children <5 years [12, 13, 33, 72, 140, 144].

Table 7: Characteristics and findings from economic evaluation studies

CHARACTERISTICS OF STUDIES							MAIN RESULTS	
STUDY ID	**Objective & setting**	**Design & parameters**	**Perspective**	**Inputs**	**Participant**	**Interventions**	**Cost-savings**	**Cost-effectiveness**
Shillcutt et al 2008	Cost saving & Cost-effectiveness of RDT vs PT in context of ACT; OPD, rural, SSA	Design: CEA using decision tree model; Parameter: age, accuracy of diag. test, malaria prevalence	Providers	RDT, anti-malarials, antibiotics, staff time, training, supervision	outpatients, febrile	RDT vs PT, context of ACT (AL)	Probability of RDT saving cost at low to moderate prevalence (0% - 40%) is 70% to 80%	ICER = cost per DALY averted < $150 95% certainty that RDT is cost-effective at prevlence <62%; 50% certainty at prevalence <81%
Zikusooka et al 2008	Cost savings due to use of RDT, in the context of 2 ACTs (AL, AS-SP), different malaria prevalence, different age sub-groups	Design: Micro-costing using population data; Parameter: age, accuracy of diag. test, malaria prevalence	Providers	RDT & antimalarials only	outpatients, febrile	RDT vs PT, in the context of AL vs AS-SP	Cost saving at prevalence up to 52% for AL; up to 29% if cheaper ACT (ASSP) used; Not cost-saving in children	
Zurovac et al 2008	Cost savings due to use of RDTs, in the context of 2 ACTs , different malaria prevalence, different age sub-groups	Design: micro-costing using patient-level data from an RCT. Parameter: age, accuracy of diag. test, malaria prevalence	Providers	RDT, ACT, antibiotics	outpatient febrile & afebrile; 5+ years	RDTvs PT	Cost reduction of up to 21% in moderate prevalence setting (26 - 38%); cost increase of 41% in low transmission setting	
Rolland et al 2006	Cost saving due to use of RDT vs PT in context of ACT; OPD, in context of epidemics	Design: decision tree modelling; Parameter: age, accuracy of diag. test, malaria prevalence	Providers	RDT, ACT, antibiotics	outpatients febrile	RDT vs PT	Cost saving at prevalence up to 55% for AL; up to 21% if cheaper ACT (ASAQ) used;	
Lubell et al 2008 (a)	Cost saving & Cost-effectiveness of RDT vs PT in context of ACT; OPD	Design: decision tree model; Parameter: age, accuracy of diag. test, malaria prevalence	Providers	RDT, ACT, antibiotics	outpatients febrile	RDT vs PT	RDT is cost-saving at low to moderate malaria prevalence (<50%)	At a DALY value of $ 150, use of RDT is preferred up to a prevalence of about 70%
Lubell et al 2008 (b)	Cost savings RDT, in the context of ACTs , different malaria prevalence, different age sub-groups	Design: decision tree-based, interactive model Parameter: age, accuracy of diag. test, malaria prevalence	Providers	RDT, ACT, antibiotics, staff time, supervision	outpatient, febrile	RDT vs PT	Cost-savings: 30% in low prevalence; 25% in moderate prevalence; not cost-saving in children	ICER = cost perDALY averted < $150 Cost-effective at low to moderate prevalence; not cost-effective at high prevalence; not cost effective in children

AL: artemether-lumefantrine; ASSP: artesunate+SP; OPD: out-patients' department; PT; presumptive treatment; SSA: sub-Saharan Africa

Furthermore, the review indicates that the degree of cost saving is dependent on the range of inputs included in the analysis. When a broader range of inputs are included (antibiotics, staff time, supervision, training), then use of RDT is associated with a smaller degree of cost reduction, or is less likely to be cost-saving [72, 140]. Finally, the studies show that use of RDTs is more cost-saving when used in the context of more expensive anti-malarials [12, 13]. Use of RDT-based guidelines could be associated with added cost to the system as the price of anti-malarials falls.

Notably, the Tanzanian trial-based study which incorporated actual data on clinician adherence shows the opposite picture. Whereas use of RDT-based guideline was associated with a 21% reduction in cost in settings with moderate malaria prevalence (26% - 38% prevalence), use of the same guideline was associated with a 41% increase in cost in settings with low malaria prevalence (0% to 1.5% prevalence) [33]. The reason for this was that health workers did not follow the RDT-based guideline in the latter setting. Most participants were prescribed anti-malarials either presumptively or despite having negative RDT results [33]. A CEA model developed by Lubell and colleagues in 2008 shows that, at the level of anti-malarial prescribing to RDT-negative patients found in observational studies and trials, use of RDT-based policy could increase healthcare cost by 10% to 250%, depending on the transmission rate [144]. Therefore, the findings of the Tanzanian trial, and of the model by Lubell et al, underline the significance of clinician adherence in all malaria transmission settings if the objectives of RDT-based guidelines have to be realised. Therefore, investment in methods to improve adherence to guidelines is essential.

Cost-effectiveness

Three studies analysed and presented cost-effectiveness of RDT-based policy as the incremental cost per DALY averted by its use, relative to presumptive treatment [72, 140, 144]. The DALY is a measure of health outcome that incorporates both premature death and morbidity or disability, including the harm of treatment. The DALY caused by a disease consists of the years of life lost (YLL) and the year of life lived with disability (YLD) [141, 143]. One DALY is valued at US$ 150. Therefore an intervention that costs <US$150 to avert 1 DALY is considered cost-effective [72, 140, 144].

When health outcomes are measured in terms of broader measures such as the DALY, use of RDT-based guideline is likely to be cost-effective (95% certainty) at malaria

prevalences of up to 62% [72, 140, 144]. That is, the cost of averting 1 DALY due to fever is likely to be < US$150 within the malaria prevalence in the 0% to 62% range. At higher prevalence levels, the incremental cost per DALY averted is likely to be >US$150 [72, 140, 144].

e) Conclusion

Use of RDT-based guideline is likely to be cost-saving and cost-effective if used in low to moderate malaria prevalence areas, as long as clinicians comply with the guidelines most of the time; and especially if RDT use is restricted to older patients. A fall in the price of anti-malarials may favour presumptive treatment. It is essential to quantify the potential clinical and economic impacts of RDT-supported policies using actual field data so that value for money can be accurately assessed.

2.9 Implementation of clinical guidelines in Africa

This section reviews literature on the quality of implementation of clinical guidelines in the African context. It shows how the components of healthcare process evaluated can affect the reported quality of implementation. The search strategy for this chapter, including the electronic databases, has been described in section 2.2. The search strategy for this section is shown in Appendix 3 (d), search set 38. The review included studies which assessed implementation of any clinical guideline in Africa between 1990 and Sept 2011.

2.9.1 Main features of studies reviewed

Table 8 below describes the relevant features of the studies included in the analysis.

Table 8: Characteristics of the studies included in the meta-analysis

N	Characteristics	No. of studies	References
1	No of studies (total)	15	
	W. Africa	6	[113-118]
	E. Africa	4	[62, 92, 111, 112]
	S. Africa	3	[91, 109, 110]
	N. Africa	1	[90]
	C. Africa	1	[108]
3	Guidelines considered by study		
	IMCI-related	8	[62, 90, 92, 109, 113, 115, 117, 118]
	IMCI + malaria	1	[112]
	Fever/malaria	4	[91, 108, 112, 146]
	Other disease-specific	1	[110]
	INRUD	1	[114]
4	Procedures assessed by study		
	Assessment only	2	[114, 118]
	Treatment only	7	[91, 108-112, 116]
	Assessment, classification and treatment	2	[62, 115]
	Assessment, classification, treatment and counselling	4	[90, 92, 113, 117]
5	Methods for data collection		
	All 15 studies used a combination of clinical observation, interview of HW and clients, and review of records (both clinical and inventory of supplies)		
6	Adjustment for clustering and weighting		
	Not indicated	4	[109, 110, 114, 118]
	In all the remaining studies, results were adjusted using different methods		

a) Number and settings

Fifteen studies were included in the analysis of adherence scores, 13 of which were from West, East, and Southern Africa.

b) Design

Majority (13/15) of the studies were observational (cross-sectional surveys). Two studies were cluster randomised trials (cRCTs) designed to evaluate effects of interventions on the quality of guidelines implementation—consisting of pre- and post-intervention assessments of performance. For these 2 cRCTs, only the pre-intervention results are synthesised in this section. The post-intervention results are synthesised in the following section that examines the effect of various interventions on the quality of guidelines implementation.

c) Types of guidelines included in studies

Nine of the studies assessed adherence to guidelines for Integrated Management of Childhood Illness (IMCI): in their original forms, or modified versions, or alongside other guidelines. Four studies assessed adherence to guidelines targeting fever/malaria management; while one study considered a guideline targeting 4 other specific conditions. One study assessed adherence to a guideline developed in the framework of International Network on Rational Use of Drugs (INRUD), covering all age groups and conditions.

d) Components of healthcare assessed by studies

Approximately half of the studies (7/15) assessed adherence in terms of treatment-diagnosis match only. Few (4/15) studies assessed all 4 components of care: assessment, classification, treatment and counselling.

2.9.2 Measure of performance

The studies reviewed present performance in form of a *score,* defined as the percent of all the expected tasks that were performed as per the guidelines under review. They provide adherence scores for individual procedures and/or for all the procedures combined. In addition, they provide the raw data used in the calculations. This review uses the raw data extracted from the reviewed studies to calculate performance scores for the components of guidelines assessed. Calculations were undertaken using MS Excel 2007. The 95% Confidence Interval (CI) around the scores were calculated using the formula for CI for single proportions [147], namely:

$$CI = p \pm z^*\sqrt{pq/n}$$

Where: z = z-score (1.96 for 95% CI)

p = performance score (percent of tasks performed as per guidelines)

q = (1-p)

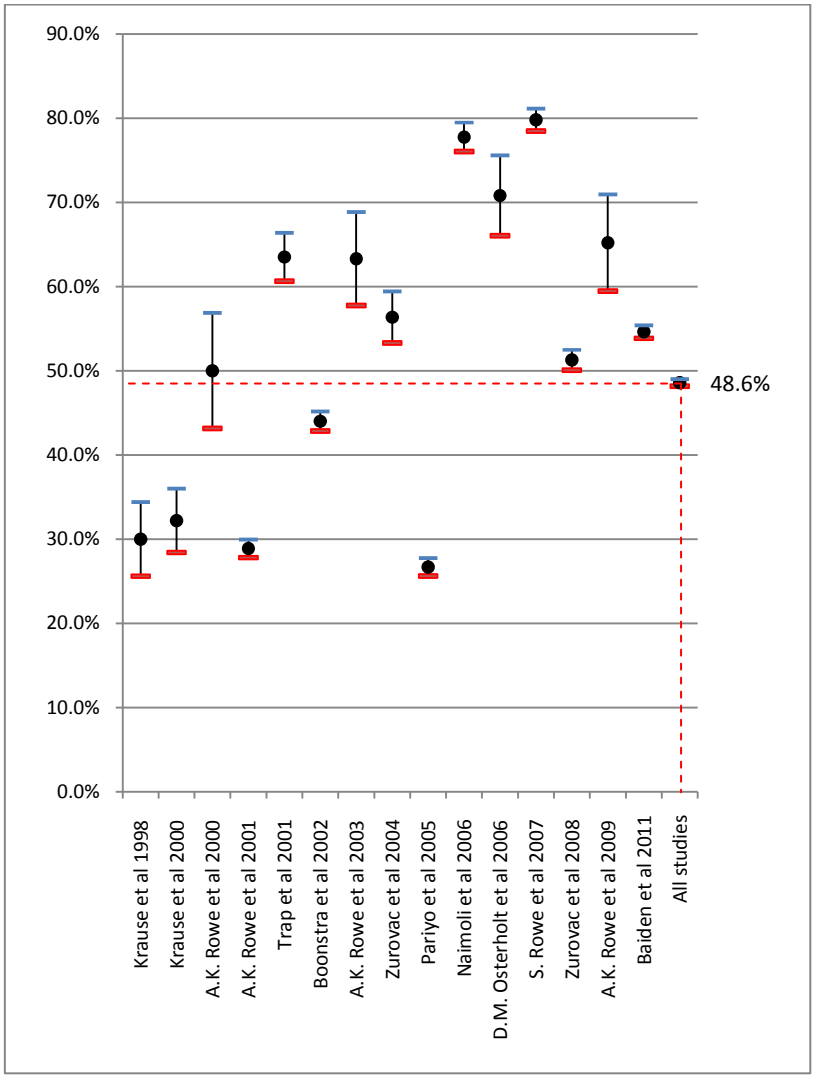
n = total number of tasks analysed

2.9.3 Main findings

a) Average scores within and across studies

A total of 52,678 tasks were analysed, out of 16,480 consultations (average of 3.2 assessment tasks per consultation). Figure 7 below shows the combined performance scores (percent of required tasks performed) for individual studies, and for all the studies combined.

Figure 7: Percent of required tasks performed by studies, with 95% confidence interval



The percent of clinical tasks performed as required by the guidelines evaluated varies widely, from 26.7% to 79.8%. The combined performance score for all the included studies is 48.6%; that is, of the 52,678 tasks analysed, 48.6% were performed according to the recommendations of the guidelines evaluated by the reviewed studies. The

variability is partly explained by the variability in the types of guidelines reviewed, and by the variability in the number and types of procedures and tasks assessed.

b) Procedures assessed and adherence by specific procedures

Table 9 below shows the different components of the healthcare process evaluated by the studies reviewed, and the combined performance scores for the procedures evaluated.

Four types of clinical procedures were examined by the 15 studies, in various combinations:

i) Assessment, comprising medical history taking and clinical examination
ii) Classification (diagnosis)
iii) Treatment (treatment-diagnosis match)
iv) Counselling (advising clients on treatment)

Table 9: Tasks performed as required by types of procedures

Procedures assessed by studies	Number of studies included in the synthesis	Number of tasks	Tasks performed as required	
			Average	(95% CI)
Assessment	6*	23,863	49.8%	(49.3% - 50.3%)
Classification	4*	2,019	38.8%	(37.8% - 41.0%)
Treatment	13	17,814	57.5%	(57.2% - 58.3%)
Counselling	4	8,987	30.3%	(29.8% - 31.2%)

** Studies that provided procedure-specific performance scores are fewer than the total number of studies that assessed the specified procedure; some studies provided only average scores for all the procedures combined.*

Most (13/15) of the studies reviewed evaluated the quality of treatment as an indicator of HW performance—either as the sole indicator of performance (7/15 studies) or alongside other components of care. Fewer studies included other aspects of care (assessment: 6 studies; classification: 4 studies; counselling: 4 studies).

The highest performance scores relate to the quality of treatment. Overall, the treatment prescribed across the 13 studies matched patient classification 57.5% of the time [62, 90-92, 108-113, 115-117]. The most problematic procedure was counselling, which was performed as expected 30.3% of the time [90, 92, 111, 113]. Patients were classified correctly 38.8% [90, 92, 114, 115], while assessment was carried out as expected 49.8% of the time [90, 92, 111, 113, 114, 118]. Since performance varies by procedures, the

average performance score reported by a study is likely to be influenced by the combination of procedures assessed.

c) Adherence by combination of procedures assessed

Table 10 below shows the various combinations of procedures included by the reviewed studies and the performance scores for the different combinations of procedures.

Table 10: Percent of tasks performed by number of procedures assessed

Procedures assessed by studies	Number of studies	Number of consultations	Number of tasks	Tasks performed as required	
				Average	(95% CI)
Assessment only	2	2296	7231	44.4%	(43.9% - 45.2%)
Treatment only	7	6402	16,454	49.9%	(49.5% - 50.7%)
Assessment, classification, treatment	2	776	10,312	45.9%	(45.5% - 46.9%)
Assessment, classification, treatment & counselling	4	7,006	9,631	39.9%	(39.4% - 40.9%)

Results from studies which evaluated the quality of treatment alone show that health workers performed tasks according to guidelines 49.9% of the time [91, 108-110, 112, 116]. However, studies which evaluated 3 or 4 components of the health care process report lower performance scores. For example, studies that included all the 4 dimensions of the health care process show that health workers performed tasks according to guidelines only 39.9% of the time [90, 92, 113, 117].

Therefore, studies that measure performance in terms of treatment-diagnosis match only are likely to report a higher quality of performance than is the case when more or all aspects of the healthcare process are assessed.

2.9.4 Conclusion

The review shows that the quality of guideline implementation varies widely in Africa. This could be due in part to the variability in the types of guidelines assessed, and to the variability in the number and types of procedures assessed by the studies reviewed. Performance scores were high for prescribing (treatment) and low for other components of guidelines. Accordingly, studies which assessed the quality of prescribing alone showed better performance scores than studies that assessed the quality of implementation of three or more components of guidelines. Assessments based on the quality of prescribing alone fail to specify the diagnostic quality of the healthcare process, which itself might be inadequate.

2.10 Effects of interventions—African context

This section presents a synthesis of evidence from Africa on the effects of various interventions on the quality of implementation of various clinical guidelines. The search strategy for this section is shown in Appendix 3 (d), search set 39. The review included all studies assessing the effects of various interventions on implementation of any clinical guideline in Africa between 1990 and Sept 2011.

Raw data were extracted from included studies and analysed using Revman 5.1 [148]. The studies analysed presented estimates of effects in the form of Odds Ratios (OR) and 95% CI. Accordingly, the results of these meta-analyses are presented in the form of Odds Ratios (OR) and 95% CI. Estimates of effects of interventions were pooled using Mantel-Haenszel method because some studies were quite small and event rates sparse. Results were analysed using random-effects analysis because of the variability in estimates of effects of the interventions. Random-effects methods assumes that variability in results is the result of systematic difference between studies and not due to random errors [99].

The review investigated consistency of effects across various studies and explored the causes of any heterogeneity in estimates of effects of interventions. Thus the review applied the Chi^2 test for heterogeneity with a 10% level of statistical significance, and the I^2 statistic with value of 0% to 40% representing an insignificant level of heterogeneity; values of >40% to 60% to denote moderate levels of heterogeneity; values >60% to 80% to represent substantial levels of heterogeneity; and values >80% to represent considerable heterogeneity—as recommended in the "Cochrane handbook for systematic reviews of interventions"[99].

2.10.1 Training

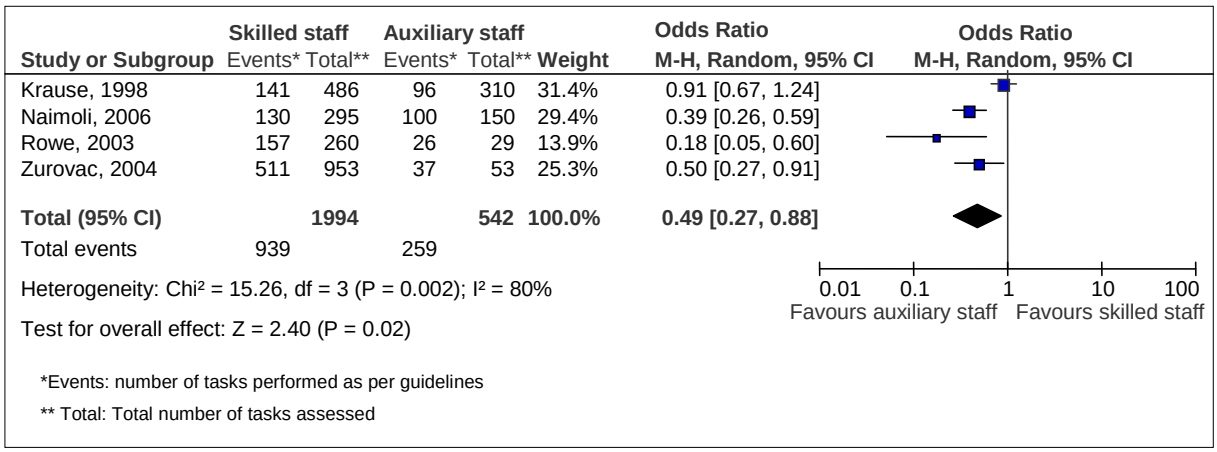
Several studies have evaluated the effects of pre-service and in-service training on the quality of guideline implementation in Africa between 1998 and 2010, the results of which are reviewed below.

a) Pre-service training (staff qualification)

Four studies evaluated the effect of pre-service training on the quality of implementation of guidelines in Africa. In this analysis, Health Worker (HW)

qualifications are used as a proxy to the types of pre-service (basic) training received by them. HW qualifications are grouped into two categories, namely "skilled staff" and "auxiliary staff". The "skilled staff" category comprised nurses/midwives, clinical officers/medical assistants, and doctors; and the "auxiliary staff" category consisted of nursing aides/nursing assistants and other paraprofessional staff generally labelled as auxiliary staff. Figure 8 below shows the number of tasks performed according to guidelines by skilled staff versus auxiliary staff.

Figure 8: Effect of staff qualification on the number of tasks performed as per guidelines



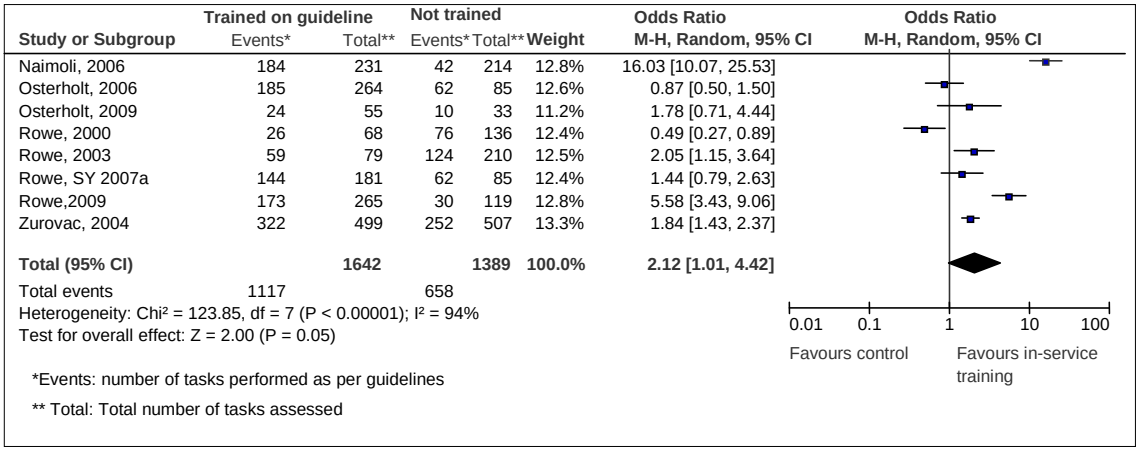
Overall, performance scores were higher amongst auxillary staff than among skilled staff. Three of the studies detected significant association between being an auxiliary staff and performing a task as recommended in the guideline under review [90, 112, 146]. Krause and colleagues did not detect a significant association between staff category and performance [114].

Implications for practice: in Uganda, RDT-based guideline has been inroduced at HCIIs, which are manned largely by paraprofessional staff (nursing assistants). The lower qualification of the majority of staff in such settings could lead to more effective implementation of RDT guidelines.

b) In-service training

Eight studies compared the performance of health workers who were trained on the use of the guidelines prior to introducing them into the health system, versus those who were not trained. The results are summarised in Figure 9 below.

Figure 9: Effect of in-service training on the number of tasks performed as per guidelines



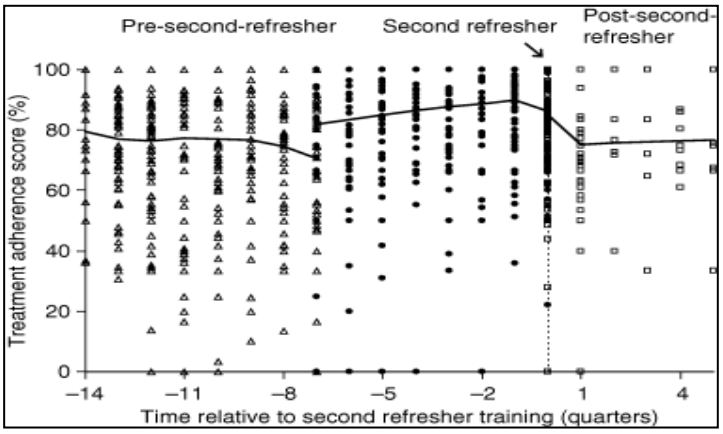
Overall, performance scores were significantly higher at facilities where the health workers were trained prior to introducing the guidelines than at facilities where they were not. There was considerable variability in effects of the intervention (I^2 = 94%, P<0.00001). Four studies detected significant association between in-service training and the number of tasks performed according to guidelines [90, 108, 117, 146]. The rest of the studies did not detect significant association between in-service training and the number of tasks performed as per guidelines [91, 93, 149, 150].

The reasons for the variability in results could be the same as given above (variability in the types of guidelines reviewed, and in the number and types of procedures and tasks assessed). The variability in results could also be explained by a possible variation in the quality of in-service trainings, and by the amount of time elapsed between the training and the studies. In most studies, assessment of implementation was done between 1 and 5 years after the initial training. Human capital theory suggests that knowledge acquired through training decays over time, especially if not backed-up by refresher training and/or constant practice [59].

c) Refresher training

Rowe and colleagues analysed the trend in clinicin adherence to IMCI guidelines in Kenya over a period of time, and the effect of refresher training on performance over time [62].

Figure 10: Treatment adherence scores over time relative to second refresher training (IMCI guideline)



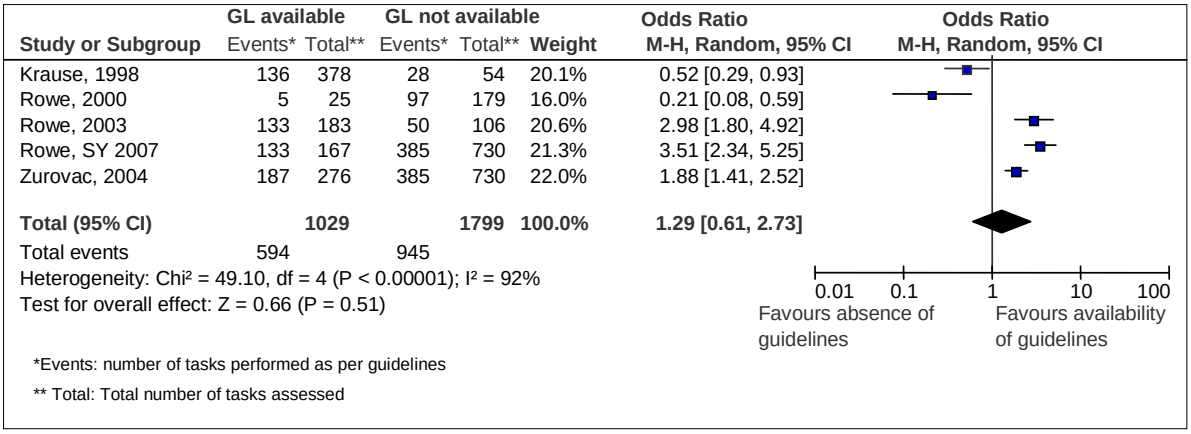
Note: quarters = quarter of a year
Source: SY Rowe (2007) [62]

The findings from their analysis (Figure 10 above) show that adherence to IMCI guideline declined steadlily following initial in-service training. The first refresher training led to improvements in adherence, above the level observed after the initial (in-service) training. However, the second refresher training was followed by a decline in treatment adherence scores. Therefore, whilst refresher training may help to reduce the rate of forgetting, the effectiveness of multiple refresher trainings is questionable. However, the results may have been due to random error or bias given that there were considerably fewer observations post second refresher training, a single study source and likely design effect.

2.10.2 Availability of guidelines

Figure 11 below shows the results from 5 studies which investigated the effect of availability of guidelines at health facilities on the quality of guideline implementation. The studies compare HW performance at facilities with guidelines on the day of the survey versus those at facilities without guidelines on the day of the survey.

Figure 11: Effect of availability of guidelines on number of tasks performed as per guidelines

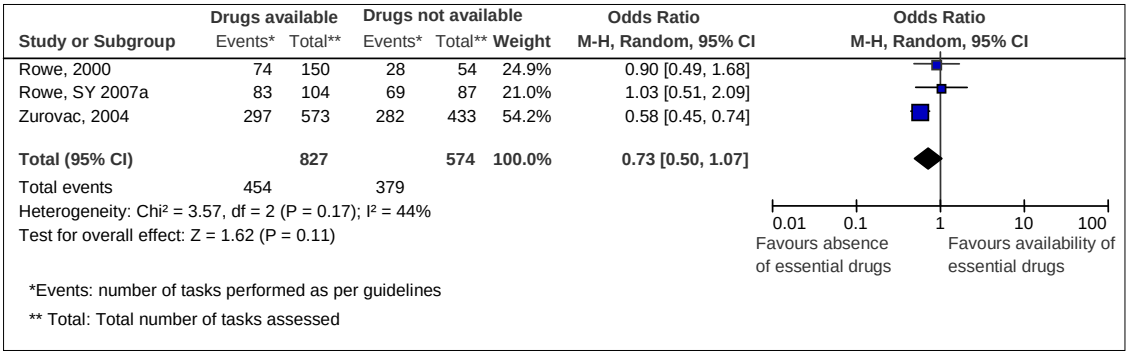


The results from the 5 studies are mixed (I^2 = 92%, P<0.00001). Two studies reveal significantly lower performance scores at facilities with job aids than at facilities without [114, 150]. Three studies show siginificantly higher performance scores at facilities with job aids than those without job aids [93, 112, 146].

2.10.3 Availability of essential drugs

Three studies investigated the association between availability of essential drugs at surveyed facilities and the quality of guideline implementation. Essential drugs were those deemed by the studies as being necessary in providing the basic health services which were specified in the guidelines assessed. They are not detailed in this review. The results from the studies are summarized in Figure 12 below.

Figure 12: Effect of availability of essential drugs on number of tasks performed as per guidelines



Two studies detected no association between availability of indicator drugs at surveyed facilities and the number of tasks performed according to guidelines [93, 150]. One study detected significantly higher performance scores at facilities where indicator drugs were not available on the survey day than at facilities where the drugs were available on

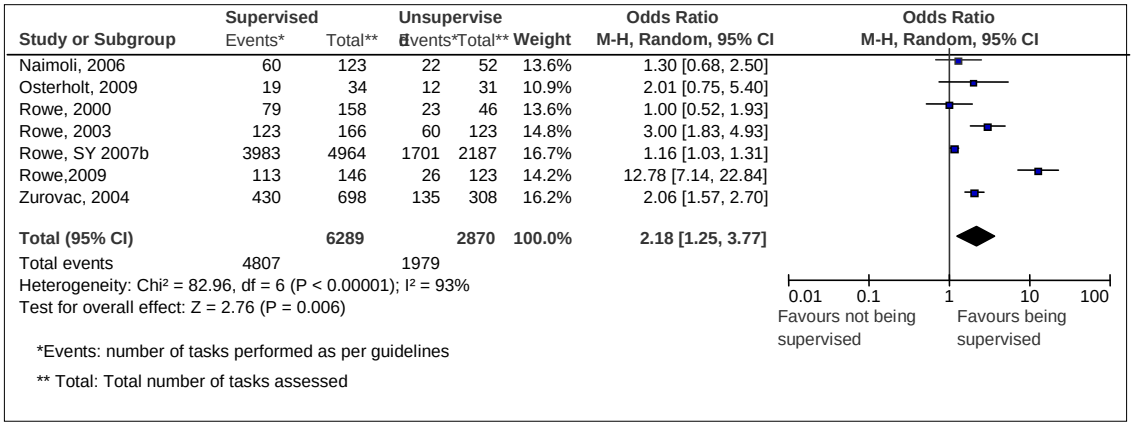
the survey day[112]. Overall, availability of indicator drugs did not seem to have an effect on the quality of guideline implementation.

2.10.4 Supervision

a) Frequency of supervision

Seven studies investigated the effect of supervision frequency on the quality of guideline implementation by comparing the performance scores for health workers who had received variable number of supervision visits over the 6 months prior to the survey. This analysis compares the performance scores for health workers supervised at least once during the 6 months prior to the survey versus the scores for HWs not supervised at all during the same period. The results are summarised in the Figure 13 below.

Figure 13: Effect of frequency of supervision on number of tasks performed as per guidelines



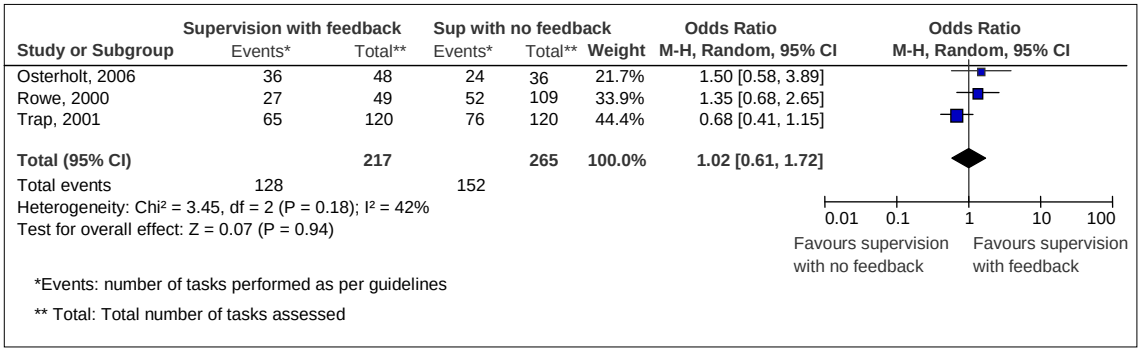
Overall, supervised health workers performed better than unsupervised health workers. However, there was considerable variability in estimates of effects of the intervention (I^2 =93%, P<0.00001). Four studies detected statistically significant difference in performance scores among the two groups of HWs, all of which were in favour of supervised health workers [93, 112, 117, 146]. However, only two of these detected a difference that was clinically significant (Rowe 2009: OR = 12.78, 95% CI: 7.14 to 22.84; Zurovac 2004: OR = 2.06, 95% CI: 1.57 to 2.70). Three studies detected no significant association between performance frequency of supervision [90, 108, 149].

b) Quality of supervision

Three studies evaluated the effect of quality of supervision on the quality of guideline implementation [91, 110, 150]. "Quality" supervision was defined as a supervisory visit

during which the supervisor observed and gave feedback on the health worker's performance. Figure 14 summarises the results from the three studies.

Figure 14: Effect of quality of supervision on number of tasks performed as per guidelines



The 3 studies did not detect a significant association between HW performance and the quality of supervision (receiving or not receiving feedback) (OR = 1.02; 95% CI: 0.61 to 1.72; N = 265; 3 studies).

2.10.5 Conclusion

Evidence from Africa suggests that a variety of interventions can improve the quality of guideline implementation. However, their effects are widely variable and may be context specific. Notably, the evidence suggests that paraprofessional staff are likely to comply with clinical guidelines better than professional staff. The evidence from this review underscores the significance of in-service training as a predictor of effective guideline implementation. However, the effect of in-service training on the quality of guideline implementation appears to be short-lived. Whilst refresher training may help to reduce the rate of forgetting, the effectiveness of multiple refresher trainings seems questionable—although these results need to be intepreted with caution as it is based on a single study source, with considerably fewer observations post second refresher training, and because of a likely design effect.

It is not clear from the available evidence if availability of guidelines at facilities has a sigificant bearing on health worker performance. However, availability of indicator drugs does not seem to have any effect on health worker adherence to guidelines. While the evidence suggests that exposure to supervision can improve the quality of guideline implementation, the quality of supervision (in terms of providing feedback to health workers) does not seem to be an important factor.

2.11 Summary of literature review

The diffusion model (2.3) specifies the type and sequence of research required in assessing the value and applicability of a new healthcare technology before dissemination for routine use. It identifies specific attributes for assessment at each level of research, which is useful in assessing complex healthcare interventions such as the RDT-based guideline. The Diffusion of innovation model is complimented by the Donabedian model of health systems performance (2.5.2) which shows the causal linkage between structural, process and outcomes attributes of quality.

Several levels and types of research are required in assessing a technology, from product development (phases 1 and 2) through to demonstration (phase 5). With regard to RDTs, extensive phase 1 and phase 2-level evaluations (product developments and the quality of manufacturing) have already taken place and are on-going (2.4). The current question is whether RDT-based guidelines can be effective in clinical settings and if they can be implemented sufficiently in clinical practice. Therefore, this thesis focuses on phase 3 through 5- level research (efficacy, effectiveness and implementation), basing on the current debates and theory.

A systematic review of RCTs was chosen to evaluate efficacy of the intervention because it was not feasible to undertake an RCT in the context of this thesis due to financial and time constraints. RCTs are considered the best design for attributing outcomes to an intervention. However, the literature suggests that RCTs are characterised by several methodological limitations that may limit the power of RCTs to detect effects of a complex intervention. In particular, the Donabedian model of health systems performance suggests that the causal linkage between a complex public health intervention such as the one under evaluation, and clinical outcomes, is weak and may be difficult to demonstrate by means of an RCT. The thesis draws from this knowledge in discussing results of chapter 3 (assessment of efficacy) and in chapter 6 (general discussion). RCTs are also critcised for the potential threats to external validity of evidence. In the context of this thesis, knowledge of efficacy is necessary to aid interpretation of evidence from effectiveness trials. Therefore, this limitation is irrelevant. A systematic review has the advantage that it can pool together a variety of evidence from a broader context. As such, it might represent the best evidence for decision making at a broader level. Literature suggests that systematic reviews can lead to generalisable conclusions by identifying similar effects in various populations.

In sub-section 2.5.5 the literature proposes a number of quasi-experimental designs which can be used in assessing effectiveness of an intervention, such as basic pretest-posttest design, control-group pretest-posttest design, and intervention versus control group comparison. The Ugandan effectiveness trial (chapter 4) used a control-group pretest-posttest design which is useful in assessing the role of maturation and history.

Demonstration studies can show the extent to which various elements of a multi-component intervention are actually implemented. In addition, it can identify key factors within the user system which can affect implementation of the policy in routine practice—which may need to be investigated, or which may need to be addressed prior to or during wide-scale implementation (2.3). For example, guideline implementation may reflect structural challenges (e.g. availability of relevant inputs, support training and supervision, policies and patients' expectations). However, according to Donabedian model of systems performance (2.5.2), compliance with structural standards does not mean that high quality care is being provided; nor does their use in quality assessment imply that high quality care cannot be provided unless these standards are complied with. Knowledge of the validity of the causal linkage between structure, process and outcomes is useful in interpreting results of chapters 3, 4 and 5.

The literature describes implicit and explicit methods used in measuring HW performance, and indicates that explicit methods such as the LQAS are much more strict and action-oriented (2.5.2). LQAS method is used in chapter 5 to assess guideline implementation. The LQAS method has several advantages over implicit methods which measure quality in terms of point estimates (e.g. percent, averages). Judgement of quality is based on explicit pre-set standards, which allows the identification of poorly performing HWs or areas that require urgent action.

In section 2.7, the literature provides an overview of several biologic malaria diagnostics that have emerged over the last 100 years. Their application in rural clinical settings has been limited, mainly by technical capacity constraints and cost. Section 2.8 describes the RDT technology, types, factors influencing performance, performance in field trials, and the WHO product testing initiative. It provides an overview of factors affecting utility of an RDT-based policy in practice and reviews economic evaluation studies of RDT-based treatment relative to symptoms-based treatment. It shows that the potential effects of RDT-based policies depends on a number of factors, including the type of

assay, its diagnostic accuracy, HW adherence, malaria endemicity, and population profiles. Several reviews indicate that RDTs may fail to achieve the desired sensitivity of 95% at parasite density of ≥ 100 parasite/100 μL of blood even under controlled trials conditions. Errors in using RDTs in routine practice could significantly exacerbate the diagnostic accuracy of RDTs. Use of RDT-based guideline is likely to be cost-saving and cost-effective if used in low to moderate malaria prevalence areas, as long as clinicians comply with the guidelines most of the time; and especially if RDT use is restricted to older patients. A fall in the price of anti-malarials may favour presumptive treatment.

The effect of various factors on the utility of RDT-based policies was explored in the analysis in chapter 4. The influence of malaria endemicity informed the selection of districts (from both high and low prevalence districts) for the survey of practice in chapter 5.

Section 2.9 reviews implementation of guidelines in African context. Several studies indicate that the quality of guideline implementation varies widely. Studies vary in the components of guidelines assessed. Studies assessing fever-oriented guidelines frequently focus on the quality of prescribing, which tends to overestimate the quality of implementation. Further, studies frequently use implicit measures of performance. Evidence from Africa also suggests that a variety of interventions can improve the quality of guideline implementation (2.10). The interventions consist of factors outlined in the diffusion of innovation and the Donabedian models and are usually assessed in surveys of fever-oriented guidelines. They have been assessed in the survey in chapter 5. Paraprofessional staff are likely to comply with clinical guidelines better than professional staff. Support services such as in-service training and supervision are generally associated with improved HW adherence to guidelines. The effect of in-service training on guideline implementation appears to be short-lived; hence the need for refresher training and experiential learning. Availability of guidelines at facilities appears to have mixed effects on guideline implementation, while availability of essential drugs may to have no significant influence on guideline implementation.

Information from sections 2.9 and 2.10 are used in discussing results in chapter 5.

Chapter 3

Systematic review of randomized controlled trials

Chapter 3 Effects of RDT-based Policy on quality of care under optimal conditions

3.1 Introduction

The purpose of this chapter is to determine if RDT-based policies can lead to better quality of patient care relative to treatment based on clinical judgement, when delivered under carefully controlled conditions. To answer this question, the chapter synthesises evidence from randomised trials which compare the effects of treatment protocols based on results of RDTs for malaria versus treatment based on clinical diagnosis of people with fever. A new technology may be ineffective in real-world settings because it is not efficacious [40]. Once a new technology is shown to be efficacious under optimal conditions, it is then useful to carry out larger more pragmatic studies to assess if it can be effective in actual practice, and to carry out health services research to evaluate the quality of implementation of the intervention in actual practice [40, 41, 50].

3.1.1 Research question

Do carefully controlled trials show that treatment policies for treating fever which are based on RDT results lead to better quality of care than treatment based on clinical judgement?

3.1.2 Aim

To review evidence from RCTs to determine if RDT-based policies for fever lead to better quality of patient care than presumptive treatment under optimal conditions

3.1.3 Objectives

1) To establish the quality of implementation of RDT-based policies under optimal conditions

2) To compare the effects of treatment policies using RDTs versus clinical judgement in treating febrile patients in malaria endemic areas

3.2 Methods

3.2.1 Analyses undertaken in this review

Table 11 provides a summary of the evidence synthesized in this chapter from the trials included in the review.

Table 11: Summary of results synthesised in this review

Variables	Analyses
Implementation	HW response to negative RDT results
Primary outcomes	Patients still unwell at day 4+ of follow-up Patients prescribed any anti-malarials
Secondary outcomes	Patients prescribed antibiotics Slide positive cases missed by RDTs Slide negative cases prescribed antimalarials

3.2.2 Criteria for considering studies for this review

Studies fulfilling the following characteristics were selected for the review:

a) **Design:** individual and cluster randomised trials

b) **Interventions**: trials comparing clinical protocols based on RDTs for fever versus protocols based on clinical diagnosis of malaria. Trials comparing RDT-based protocols versus microscopy-based protocols were excluded, except if they included clinical diagnosis as a second comparison.

c) **Participants:** trials comparing the 2 protocols (in (b) above) in outpatients with fever, or with a history of fever in the preceding 48 hours; or outpatients suspected to be having malaria by clerking clinicians. The review focuses on malaria in endemic areas. Trials conducted in non-endemic areas (e.g. fever in travellers in Europe) were excluded.

d) **Outcomes:** trials evaluating any of the following outcomes:

Primary outcomes

1. Patients still unwell at day 4+ of follow-up
2. Patients prescribed any anti-malarials

Secondary outcomes

1. Patients prescribed antibiotics
2. Microscopy-negative patients prescribed anti-malarials
3. Microscopy-positive patients not prescribed anti-malarials

3.2.3 Search methods for identification of studies

We attempted to identify all relevant studies regardless of language or publication status (i.e. published, unpublished, in press, and ongoing).

a) Electronic searches

We searched the following databases on 26 January 2011 and on 28 October 2011, using the search term described in Appendix 4: (a) Cochrane Infectious Disease Group Specialized Register, (b) Cochrane Central Register of Controlled Trials (CENTRAL), published in *The Cochrane Library* (2011 issues 1&4), (c) MEDLINE, (d) EMBASE, (e) CINHAL, (f) PschINFO, and (g) Science Citation Index.

In addition we searched the metaRegister of Controlled Trials (mRCT) and the WHO trials register using "malaria" AND "rapid diagnostic test*" OR "presumptive treatment" as search terms *(*26 January 2011 & 28 October 2011).

This search strategy yielded a total 233 abstracts of trial reports, after removing duplicates (first search: 202; repeat search: 31); and 8 records of on-going trials. Table 12 summarises the sources searched and the hits retrieved from each.

Table 12: Electronic databases searched and hits retrieved

	Hits Retrieved		
SOURCE	January 2011	October 2011	TOTALS
REPORTS			
Cochrane Infectious Disease Group Specialized Register	21	13	34
MEDLINE (PubMed)	239	64	304
EMBASE	59	10	69
Cochrane CENTRAL	150	3	153
Science Citation Index	116	24	140
PsycINFO	0	1	1
CINHAL	5	9	14
Final number of records in database			
before deleting duplicates & irrelevant titles	**590**	**124**	**715**
after deleting duplicates and irrelevant titles	**202**	**31**	**233**
ONGOING STUDIES			
WHO Trials register	1	0	1
mRCT website	6	1	7
Final number of ongoing trials after deleting duplicates	**7**	**1**	**8**
ALL MATERIALS			
before deleting duplicates & irrelevant titles	**597**	**125**	**723**
after deleting duplicates and irrelevant titles	**209**	**32**	**241**

b) Searching other resources

Researchers and Organization

In addition, we contacted researchers in the field to identify additional studies that may be eligible for inclusion. This strategy did not yield additional materials.

Reference lists

Furthermore, we checked the reference lists of all trials selected from studies identified by the search strategy described above. This strategy also did not yield additional materials.

3.2.4 Data collection and analysis

i) Selection of studies

John Odaga (JO) and Joseph A. Lokong (JAL) independently screened the abstracts in the search list generated by the search of electronic databases for potentially relevant articles. We applied 2 criteria to identify potential candidates for inclusion in the review. An abstract was listed for further scrutiny if it satisfied both of the following selection criteria:

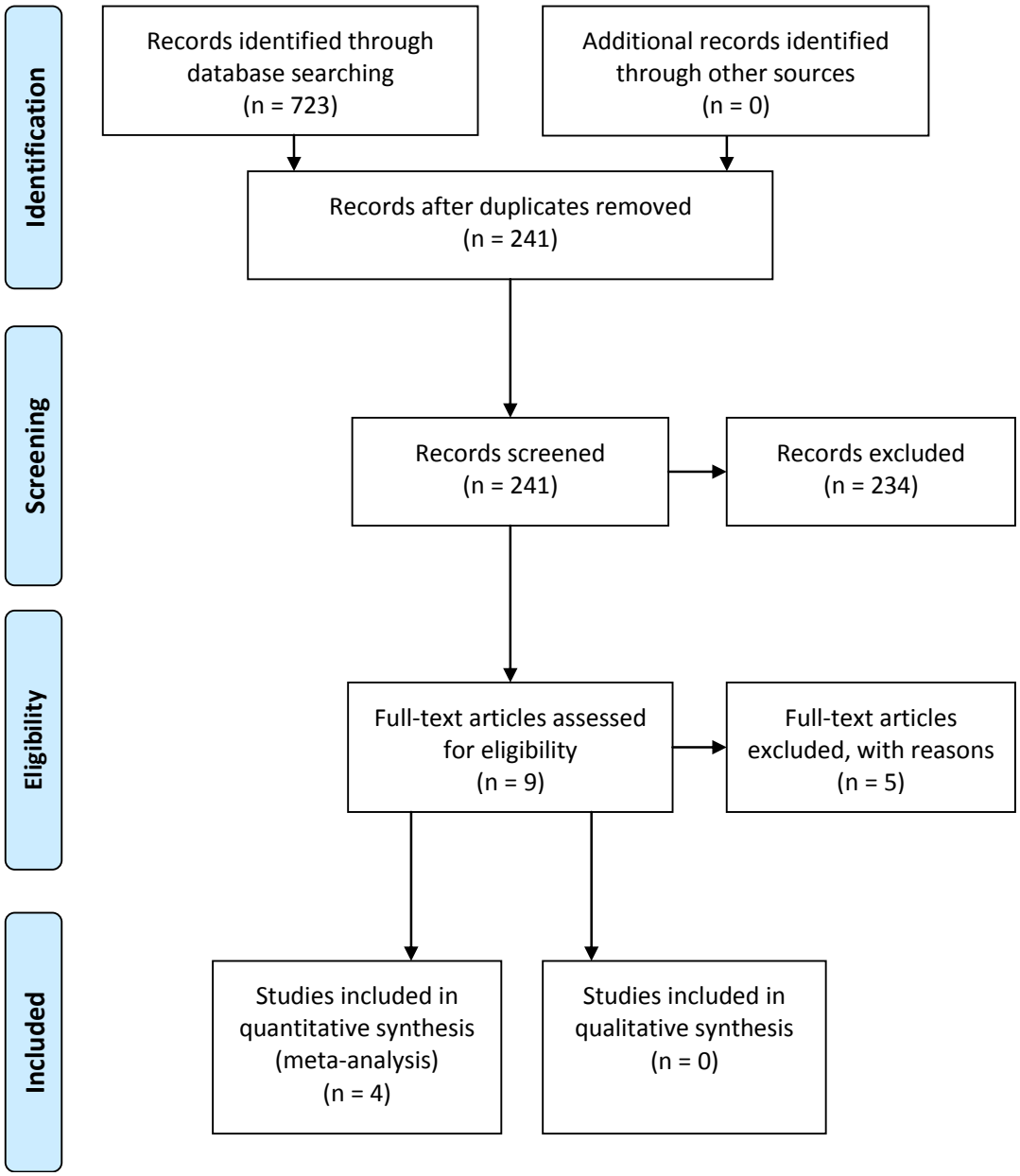
a. *interventions compared*: if one of the comparison interventions in the trial was a policy based on malaria RDT
b. *outcome:* if the trial evaluated at least one of the outcomes listed in 3.2.1 (d) above. Trials which evaluated accuracy of RDTs, but which did not evaluate their effects on the outcomes listed above were dropped.

The selection process is illustrated in the flow diagram (Figure 15) below.

JO and JAL compared their lists of potentially relevant titles. Both authors identified the same studies (9) for possible inclusion in the review. JO retrieved the full texts of the selected (9) articles, which were made available to both authors. Both JO and JAL independently assessed each of the 9 studies to select those to include in the review, based on the inclusion criteria listed in 3.2.1(a – d) above. A study was included in the review if it satisfied all of the characteristics described in 3.2.1(a – d) above (study types, settings, population, interventions, comparison and outcomes). JO and JAL discussed the lists of studies identified for inclusion between them. Any disagreements were resolved by referring to the original articles and/or through discussions, and, where

necessary, by consulting Paul Garner (PG) and Sarah Donegan (SD). Four trials were included in the review [34-37], and are described further in Characteristics of included studies table Appendix 5. Trials that were excluded are listed in the Characteristics of excluded studies table, which also describes the reasons for the exclusion Appendix 6 [32, 120, 121, 151, 152]

Figure 15: Diagram showing electronic records identified, screened and included in synthesis



Source of diagram: Moher D, Liberati A, Tetzlaff J, Altman DG, The PRISMA Group (2009). *P*referred *R*eporting *I*tems for *S*ystematic Reviews and *M*eta-*A*nalyses: The PRISMA Statement. PLoS Med 6(6): e1000097. doi:10.1371/journal.pmed1

(www.prisma-statement.org)

ii) Data extraction and management

JO and PG independently extracted outcomes data from the included trials, guided by a standard data extraction form (Appendix 7). Any disagreements were resolved by referring to the original paper and through discussions. Where necessary, clarification was sought from trial authors. Study authors were directly contacted to provide relevant data found missing in the included studies (e.g. number of participants by age group, the number of health centres excluded from analysis).

Extracting data from cluster randomised trials

Two of the included studies were cluster randomised trials in which the unit of randomisation were health facilities but analyses performed at patient level [35, 37]. The ideal data to extract in this circumstance is the direct estimate of the measure of effect (e.g. RR with its CI) from an analysis that accounts for clustering [99]. The 2 cluster randomised trials included in this review [35, 37] differed in the extent to which they dealt with clustering with regard to various outcomes of interest.

Yeboah-Antwi and colleagues (2010): JO and PG extracted estimates of measures of outcome which were adjusted for clustering & baseline imbalance using Generalised Estimating Equations (GEE) with Exchangeable Correlations Matrix (ECM) [37]. JO entered extracted data (adjusted RR and 95% CI) into RevMan 5.1[148] using the Generic Inverse Variance (GIV) method as recommended in the "Cochrane handbook for systematic reviews of interventions" [99]. GIV analysis is a statistical method of combining estimates of effects calculated from raw data with estimates of effects which have been extracted from studies and entered into Revman 5.1 as such (e.g. as RR & 95% CI rather than as raw data) [99].

Skarbinski and colleagues (2009): Skarbinski and colleagues [35] took into account a design effect of 2 in their sampling design. The reported measure of outcome—net change in prescribing of antimalarials—was not appropriate for this review. Therefore, JO and PG independently extracted summary data from which JO calculated crude RR and 95% CI using RevMan 5.1[148]. JO adjusted the 95% of the RR for clustering using the approximate analysis method which involves inflating the standard errors using a design effect of 2—as recommended in the "Cochrane handbook for systematic reviews of

interventions"[99]. JO entered the adjusted estimates into Revman 5.1 [148] using the GIV method[99].

Extracting data from individual randomised trials

Two studies were individual randomised trials [34, 36].

Bissoffi and colleagues (2009): The review authors independently extracted the direct estimates of measures of outcomes (RR and 95% CI), and number of participants. JO entered data into RevMan 5.1 [148]using the generic inverse variance method [99].

Ansah and colleagues (2010): Effects of the intervention were captured in the form of ORs and 95% CI. These estimates were left in crude form for some outcomes and adjusted for clustering and potential confounders for others. The authors independently extracted raw data outcomes from which JO calculated RRs and 95% CIs for all relevant outcomes. JO entered estimates of measures of outcome into RevMan 5.1 [148]using the generic inverse variance method [99].

iii) Assessment of risk of bias in included studies

JO and PG independently assessed and judged the quality of the selected papers using the criteria described in the Cochrane's collaboration's tools for assessing risk of bias [99]. Risk of bias was assessed against seven items: (a) how allocation sequence was generated (b) how allocation was concealed to participants, investigators and outcome assessors; (c) blinding of participants and investigators; (d) blinding of outcome assessors; (e) completeness of outcomes data (number analysed relative to number randomised) (f) selective reporting: whether all pre-specified outcomes are reported; (g) other sources of bias. The review authors judged and classified the degree of risk of bias in each study along each risk of bias item as "High Risk", "Low Risk", or "Unclear" [99]. In addition the review authors summarised the degree of a particular risk of bias across all included studies [99].

iv) Measures of guideline implementation

Because of data limitation, only 2 indicators are used to assess HW performance; namely, the proportion of RDT-positive patients who are prescribed antimalarials, and RDT-negative patients for whom antimalarials were prescribed. HWs were expected to respond correctly—prescribe antimalarials to RDT-positive patients and to withhold antimalarials to negative RDT patients—at least 95% of the time. Otherwise

implementation of the guideline or HW performance was judged to be inadequate [104, 106, 127, 128, 140]. Analysis was undertaken in MS Excel 2007.

v) Measures of treatment effect

Estimates of effects of the intervention were measured in the form of Risk Ratio (RR) and 95% CI. To permit meta-analysis, data were entered into RevMan 5.1 using the GIV approach—by entering the natural logs of treatment effects and the natural logs of their standard errors [99, 148].

vi) Unit of analysis issues

Analyses of all outcomes were carried out at individual (patient) levels using generic inverse variance method, taking into account clustering and baseline imbalance as described in section 3.2.4 above.

vii) Assessment of heterogeneity

We assessed the results of the review for heterogeneity among studies by inspecting the forest plots. We applied the Chi^2 test for heterogeneity with a 10% level of statistical significance, and the I^2 statistic with value of 0% to 40% representing an insignificant level of heterogeneity; values of >40% to 60% to denote moderate levels of heterogeneity; values >60% to 80% to represent substantial levels of heterogeneity; and values >80% to represent considerable heterogeneity [99].

viii) Data synthesis

We carried out a full analysis of all included studies, irrespective of the risk of bias. Data from the studies were combined using the generic inverse variance methods in Revman 5.1 [148]. Estimates of treatment effects and corresponding standard errors were entered into RevMan 5.1 [148] using the GIV approach—by entering the natural logs of treatment effects and the natural logs of their standard errors [99]. Because of considerable heterogeneity in estimates of one of the primary outcomes (prescribing of antimalarials), data were analysed using random-effects methods. Random-effects methods incorporate the assumption that the studies are estimating different but related intervention effects. That is, the method assumes that variability in results is the result of systematic difference between studies and not due to random errors [99].

ix) Subgroup analysis and investigation of heterogeneity

We performed subgroup analysis for the primary outcomes with considerable heterogeneity [35-37] by stratifying results age-groups and degree of HW compliance with guidelines. We applied the Chi^2 test for sub-group difference with a 10% level of

statistical significance, and the I^2 statistic with value of 0% to 40% representing an insignificant level of heterogeneity; values of >40% to 60% to denote moderate levels of heterogeneity; values >60% to 80% to represent substantial levels of heterogeneity; and values >80% to represent considerable heterogeneity [99]. Data did not permit sub-group analysis by the following potential causes of heterogeneity: malaria transmission intensity, referral level of the health facility, and qualification/experience of clinicians (for anti-malarial and antibiotics prescribing). Only one study provided results by seasons [35]; hence we did not perform subgroup analysis of the combined results by season.

3.3 Results

3.3.1 Description of studies

Detailed descriptions of the studies included in the review are found in *Characteristics of included studies* table (Appendix 5). The following sections describe the main features of the studies which were included in the review.

a) Results of the search

Of the 233 abstracts and 8 titles of on-going studies retrieved, 9 studies were considered; all of which are from Africa.

b) Included studies

Four randomised controlled trials were included: from West Africa (2), East Africa (1) and Southern Africa (1) (Appendix 5)

Settings

All studies were carried out in basic healthcare facilities without microscopes, and manned mainly by paraprofessional staff. One trial consisted of 2 comparison arms: RDT-based guideline compared with microscopy-based guideline; and RDT-based guideline compared with clinical diagnosis. Only the latter arm is included in this review [36]. Two studies do not describe the malaria transmission rates in the study areas [34, 36]. Two studies [35, 37] are carried out in high and low endemicity areas but they present the results in combined form.

Design

Two of the studies randomised clusters (health facilities) [35, 37] while two randomised individual patients [34, 36]. One of the cluster randomised trials [35] incorporated parallel pre-test and post test comparisons.

Participants

Table 13 summarises the number of participants randomised and the number analysed for primary outcomes by the studies reviewed.

The number of participants randomised in the included studies is 9545. The individual randomised trials included patients of all age groups. Bisoffi and colleagues [34] included 2169 patients from 10 health centres; and Ansah and colleagues randomised 3452 patients from 3 health centres [36].

Table 13: Number of participants randomised and analysed by studies

	Ansah (2010)	Bisoffi (2009)	Yeboah-Antwi (2010)	Skarbinski (2009)	Total
Number of health facilities	3	10	31	30	74
Target population*	All	All	<5 years	≥5 years	-
Number randomised	3452	2169	3125	799	9545
Number analysed for primary outcome	3442[1]	2169[1,2]	3125[3,4]	669[1]	9405
Loss to follow-up	0.3%[1]	0.0%[1,2]	0.0%[3,4]	16.3%[1]	1.5%

**of those with fever or suspected of having malaria*

Footnotes

[1] numbers presented refer to patients prescribed antimalarials

[2]number analysed in assessing clinical outcomes was lower in Bisoffi et al: 2095; loss to follow-up = 3.4%

[3] numbers presented refer to patients assessed for clinical outcomes

[4]number of patients prescribed antimalarials in Yeboah-Antwi et al was lower: 3047; loss to follow-up: 2.5%

The cluster randomised trials targeted different age groups. Yeboah-Antwi and colleagues randomised a total 3125 children (< 5 years of age) from 31 aid posts [37] while Skarbinski and colleagues randomised 799 older patients (5+ years) from 30 health facilities [35]. See Characteristics of included studies table (Appendix 5) for further details.

Number of facilities enrolled

Studies randomising individual patients enrolled relatively fewer health facilities (3 health centres in Ghana [36] and 10 health centres in Burkina Faso [34]) than those that randomised facilities (clusters): 30 and 31 in Zambia and Kenya respectively [35, 37].

Prescribers

Table 14 shows the category and number of HWs who provided care to patients during the trials.

Table 14: Quaification and number of HWs who were assessed

Study ID	**Qualification of attending HW**	**Number of HWs/facility (average)**
Ansah, 2010-GHA	Mostly enrolled nurses, some registered nurses/midwives, a few auxiliary nurses	Number not indicated
Bisoffi, 2009-B'FASO	Nurses (level not specified)	Number not indicated
Skarbinski 2009-KEN	Not described	Number not indicated
Yeboa-Antwi 2010-ZAM	Community health workers	1.2

In 3 studies, the HWs consisted of either nurses or auxiliary HWs [34, 36, 37]. Skarbinski and colleagues did not describe the type of HWs involved in the study [35]. However, the facilities enrolled in the study included health centres and hospitals which presumably had more qualified staff (e.g. clinical officers).

Only one trial indicated the number of HWs included in the study, whereby there was mostly 1 HW per community health post [37].

Interventions

Introduction of RDT-based policies was preceded by training of health facility staff. In each of the trial from Zambia, Ghana and Burkina Faso, training lasted for 3 days [34, 36, 37]. In Kenya, HWs received training for only half of a day [35]. All trainings emphasized the use of RDTs and restricting prescribing of antimalarials to RDT positive cases only. There was no report as to whether other components of the guideline (e.g. clinical assessment, management of patients with negative RDT results) were equally emphasized. In Ghana and Burkina Faso, members of the research team carried out the tests and then sent the results to the health workers for interpretation and treatment. This intervention was aimed at optimising the quality of RDT tests and minimising work pressure on HWs. In Zambia, the community health workers (CHW) received intensive monitoring characterised by monthly assessment of HW performance and feedback. In addition, the research team provided the CHWs with bicycles to aid communication with the monitoring team based at designated health centres. No additional interventions are reported in the Kenyan trial.

Outcomes

Table 15 below shows the outcomes evaluated by the various studies included in the review. All the trials included in the review (4) report on effects of RDT-based guidelines on prescribing of antimalarials. Only one of these provide further data on effects of the guidelines on prescribing of antimalarials to slide negative patients; and on the number of missed malaria cases [36].

Table 15: Outcomes evaluated by studies included in the review

	OUTCOMES OF INTEREST			
STUDY ID*	Clinical outcomes	Prescribing of anti-malarials	Prescribing of antibiotics	Total number of outcomes evaluated
Ansah, 2010-GHA		√	√	2
Bisoffi, 2009-B'FASO	√	√	√	3
Skarbinski 2009- KEN		√		1
Yeboa-Antwi 2010-ZAM	√	√		2
TOTAL (STUDIES)	2	4	2	

**GHA: Ghana; B'FASO: Burkina Faso: KEN: Kenya; ZAM: Zambia*

Two trials report on effects of the intervention on clinical outcomes [34, 37]; and 2 report on effects of the intervention on prescribing of antibiotics [34, 36].

c) Excluded studies

Five studies were excluded and the reasons for the exclusion are summarised in the *Characteristics of excluded studies* table (Appendix 6). Studies were excluded because the designs and/or comparisons were inappropriate [32, 120, 121], or because there was no comparison [151, 152].

3.3.2 Risk of bias in included studies

This section provides an overview assessment of Risk of Bias (ROB) in the studies included in the review. Details of ROB items assessed per study and the judgements made are presented in Figure 16 and Figure 17. Detailed descriptions of each ROB item and the reasons for the judgement can be found in "Characteristics of included studies table" in Appendix 5. Further details are below.

Figure 16: Review authors' judgements about the amount of ROB per item per study

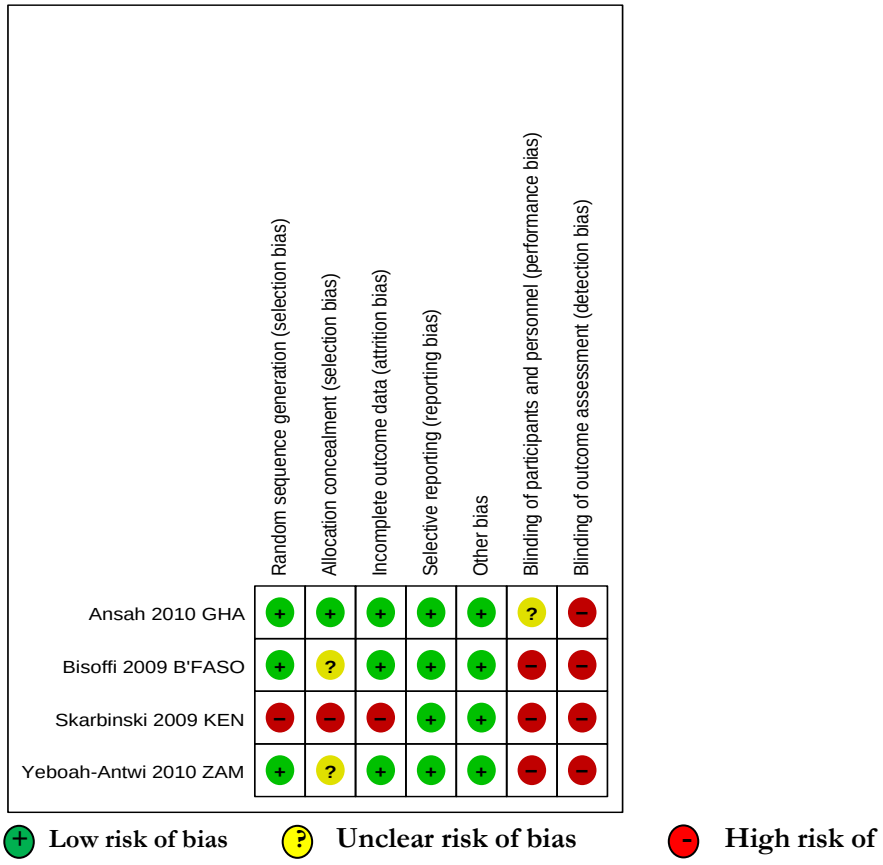
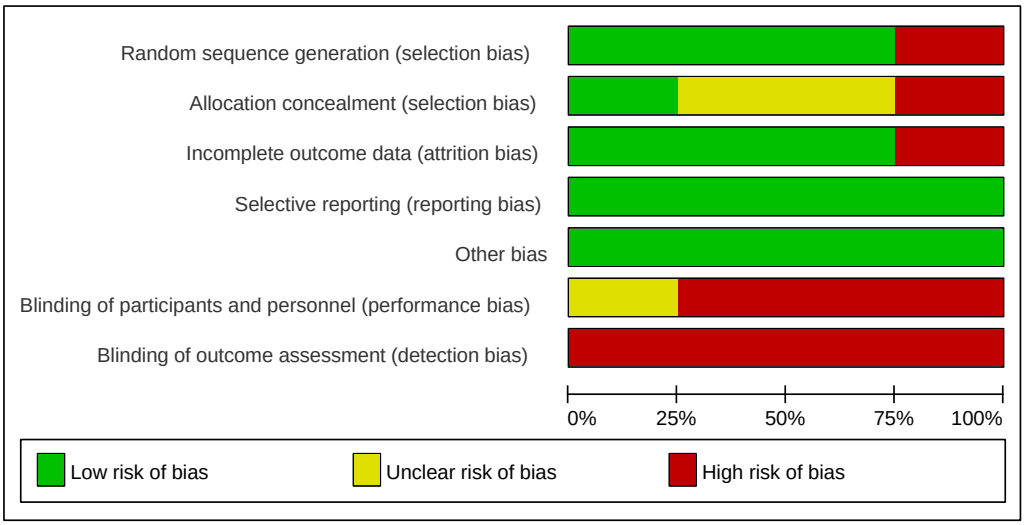


Figure 17: Cumulative percent of studies with varying levels of ROB per item



For a particular risk of bias item, percentages indicate the cumulative proportions of studies characterised by increasing levels of risk of bias, specified by the successive colour codes

In 3 trials [34-36], random numbers were generated by means of a computer. One study did not describe the method of sequence generation [37]. Ansah and colleagues

concealed allocation by placing the numbers in sealed envelopes [36]. Bisoffi and colleagues and Yeboah and colleagues did not indicate if allocation was concealed or not [34, 37]. Allocation was not concealed by Skarbinski and colleagues [35].

In all studies, participants and personnel were not blinded to the interventions, test results/diagnoses, treatment prescribed, and where applicable, treatment outcomes. No statistically significant imbalance in baseline characteristics of enrolled patients was detected between comparison groups in 3 trials [34-36]. Yeboah-Antwi and colleagues [37] report an imbalance in the two arms in the number of children presenting with fast breathing, for which adjustment was made.

a) Allocation (selection bias)

The review judged the risk of bias in sequence generation as low in 3 studies [34, 36, 37] and as unclear in 1 study [35]. Concealment of allocation was judged as low risk in one study [36], as unclear (not described) in one study [34], and as high-risk in 2 the studies [35, 37].

b) Blinding (performance bias and detection bias)

Blinding was judged as high risk in all the included studies. Given the nature of the intervention (use of a diagnostic in a clinical setting), it was practically not possible to blind the study participants and personnel of the interventions, test results, prescriptions and clinical outcomes.

c) Incomplete outcome data (attrition bias)

The review judged attrition bias as low risk (≤ 3%) in 3 studies [34, 36, 37] and as high risk in 1 study [35]. Analyses were based on available cases with complete data (available case analysis) [99].

d) Selective reporting (reporting bias)

All outcomes which were pre-specified in the methods section of the reports (and in the protocols) were reported in the results section. The risk of reporting bias was judged to be low.

e) Other potential sources of bias

Three of the studies report imbalance in the number and qualifications of the clinicians which they adjusted for using different methods [34, 36, 37]. In one study, attrition rate was high (30%) and unequal in both arms [35]. This could be a potential source of bias.

3.3.3 Quality of implementation

Two indicators were used to assess the quality of practice in a controlled condition; namely, (a) proportion of RDT-positive patients for whom antimalarials were prescribed (b) the proportion of RDT-negative patients for whom antimalarials were prescribed. HWs were expected to respond correctly to negative RDT results at least 95% of the time. Otherwise implementation of the guideline or HW performance was judged to be inadequate [104, 106, 127, 128, 140]. Only 3 studies provided data for the first indicator [34-36]. HW response was adequate in all three trials: 98% to 100% of cases with positive RDT results were prescribed antimalarials. HW performance with regard to the second indicator is summarised in Table 16.

Table 16: HW performance based on patients with negative RDTs who received antimalarials in RDT arms

	Cases with negative RDTs prescribed antimalarials			Classification of HW adherence*
Study ID	**N**	**%**	**(95%CI)**	
Yeboah-Antwi 2010 ZAM	704	0.4%	(0.3%, 0.5%)	High
Skarbinski 2009 KEN	346	41.0%	(38.4%, 43.6%)	Low
Ansah 2010 GHA	1013	49.5%	(47.9%, 51.1%)	Low
Bisoffi 2009 B'Faso	494	81.0%	(79.2%, 82.8%)	Low
All	8893	11.8%	(11.4%, 12.1%)	Low

* This classification is based on a threshold of 5% used in several health facility surveys, whereby a HW is expected to perform a given clinical task using the wrong technique less than 5% of the time; otherwise his performance (adherence) is judged to be inadequate or low [104, 106, 127, 128]; and on a model developed by Lubell and colleagues which suggests that the economic advantage (cost-saving) gained by using HRP-2-based RDTs over presumptive treatment is lost once the number of RDT negatives cases who are prescribed antimalarials exceeds 5% in children (<5 years) and/or in high prevalence areas [140].

It is notable from Table 16 that HW adherence to negative RDT results was generally low and varied widely between trials. HW adherence was high in Zambia where HWs prescribed antimalarials to only 0.4% of children with negative RDT results [37]. HW adherence was low in 3 trials [34-36]. HW adherence was notably very low adherence in Burkina Faso where HWs prescribed antimalarials to as many as 81% of children and adults with negative RDT results [34]. HWs in the 4 trials received different types of extra interventions aimed at optimising performance (see 'interventions' on page 45 above).

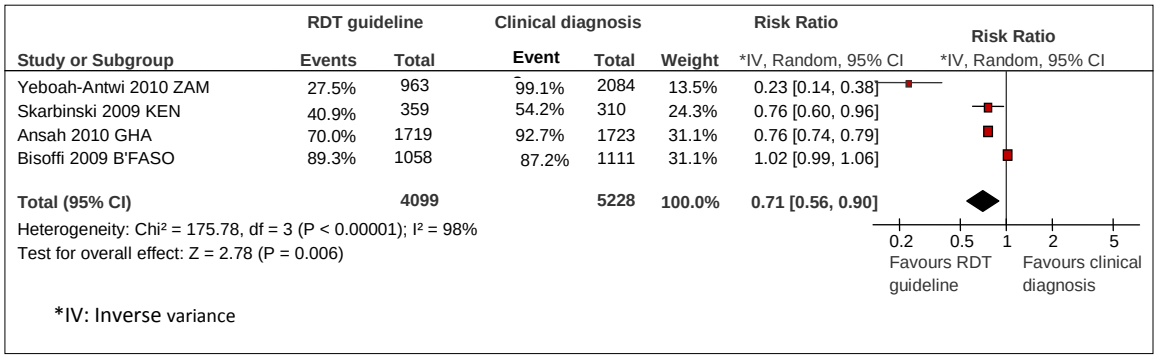
3.3.4 Effects of interventions

Primary Outcomes

a) Patients prescribed anti-malarial drugs

Overall, the four trials included in this analysis show that fewer patients were prescribed antimalarials in the RDT group (combined RR = 0.71, 95% CI 0.56 to 0.90). See Figure 18 for details. Three of the four trials detected significant reductions in prescribing of antimalarials in favour of RDT use [35-37]. One study detected no statistical difference in the proportion of patients receiving antimalarials in the two arms (RR = 1.02, 95% CI 0.99 to 1.06) [34].

Figure 18: Number prescribed anti-malarials in RDT arm versus clinical diagnosis arm

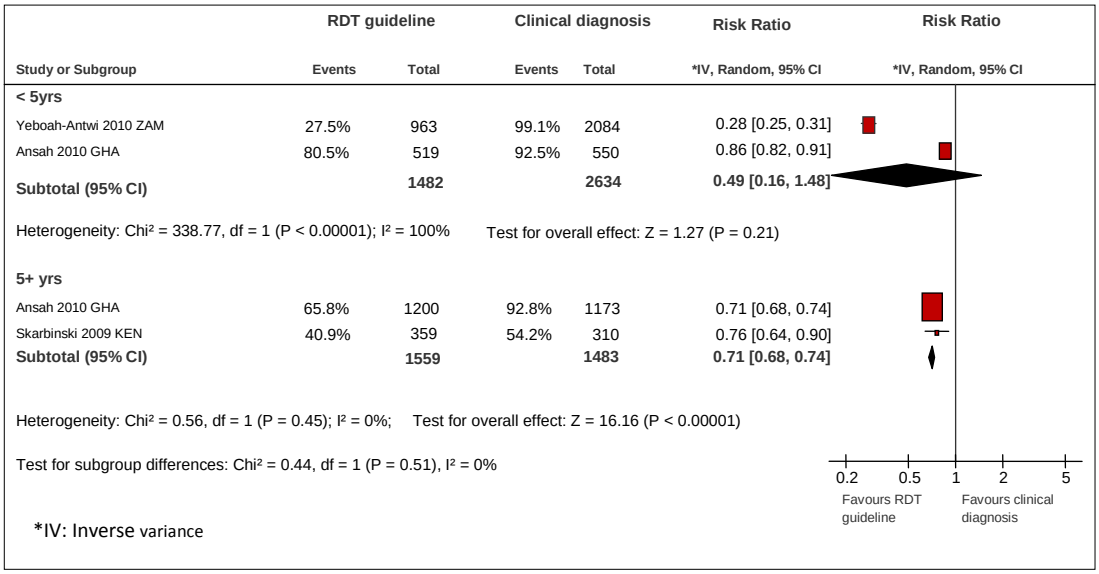


The results show considerable heterogeneity in the effects of the intervention (Chi^2 = 175.78, df = 3, (P < 0.00001), I^2 = 98%). These estimates of effects were assumed to be systematically different in the different studies for reasons specified in the literature review, section 2.8.3 (Figure 6), namely age profile of target population, HW adherence to RDT-based guidelines, and local malaria prevalence. Accordingly, sub-analyses were undertaken to investigate if the heterogeneity could be explained by variability in participants' age groups and HW adherence to RDT results. Because of data limitation, sub-group analysis was not performed to investigate the effect of other potential causes of heterogeneity.

Sub-group analysis by age group

Figure 19 displays analysis of the results by age groups (< 5 years versus ≥5years).

Figure 19: Number prescribed anti-malarials in RDT arm versus clinical diagnosis arm by age groups



The trial by Bisoffi and colleagues was not included in this analysis because data were not disaggregated by age groups.

The sub-group analysis shows the effect of the intervention was similar in both age-groups (children <5 years: RR = 0.49, 95% CI: 0.16 to 1.48, versus older children and adults 5+ years: RR = 0.71, 95%CI: 0.68 to 0.74). The test for sub-group difference shows that variability in age-group does not explain heterogeneity in estimates of effects detected between the trials (Chi^2 = 0.44, df = 1 (P = 0.51), I^2 = 0%). Therefore, in this synthesis, age was not a significant factor in the variability of estimates of effects of the intervention. It can also be noted that considerable heterogeneity in estimates of effects persists in <5 years sub-group after stratifying results by age group (Chi^2 = 338.7, df (P = 0.00001); I^2 = 100%). This further indicates that the variability in effects of the intervention may be due to factors other than variability in age-groups.

Sub-group analysis by HW adherence

Estimates of effects were stratified by degree of HW adherence to RDT results to explore if HW adherence could explain the heterogeneity in results between studies. The results are shown in Table 17

Table 17: Effect of intervention stratified by HW adherence to guideline

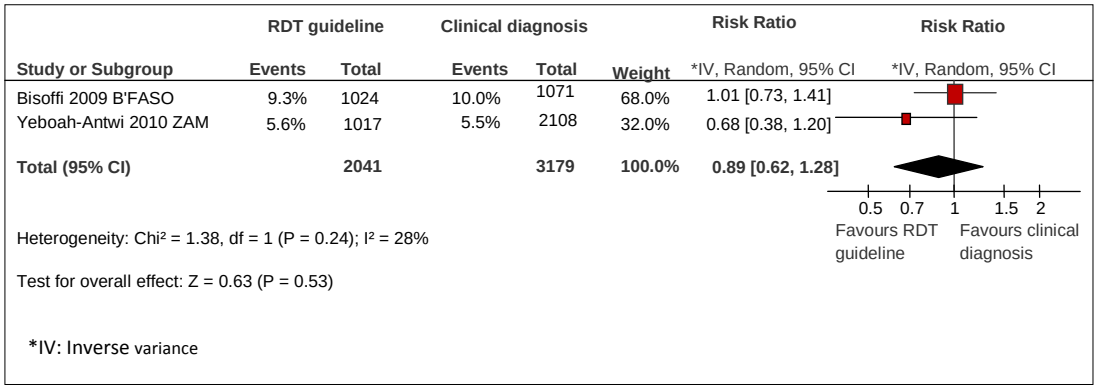
HW adherence	Studies	Total		Effect		Heterogeneity
		RDT	Clinical	RR	(95% CI)	
High	1[37]	963	2084	0.23	(0.14, 0.38)	Not applicable
Low	3[34-36]	3136	3144	0.85	(0.66, 1.08)	Chi^2 = 148.3, df = 2 (P = 0.00001), I^2 = 95.3%
Test for subgroup difference: Chi^2 = 21.44, df = 1 (P = 0.00001), I^2 = 95.3%						

Table 17 shows that systematic differences in HW adherence between the studies accounts for 95.3% of the variability in estimates of effects between the studies (Test for sub-group differences: Chi^2 = 175.78, df = 2 (P<0.00001), I^2 = 95.3%). The magnitude of effect of the intervention varied according to the degree of HW adherence to RDT results. Significant decline in prescribing of antimalarials was detected in intervention HCs where HW adherence was high (RR, 0.23, 95% CI 0.14, 0.38) [37]. No significant difference was detected in prescribing of antimalarials by the trials characterised by low HW adherence at intervention HCs (combined RR = 0.85, 95% CI: 0.66, 1.08) [34-36].

b) Patients still unwell at follow-up

No difference was detected in the proportion of patients who still had fever at follow up (seen 4 to 7 days after treatment): combined RR = 0.89; 95% CI = 0.62 to 1.28; 2 studies [34, 37]. See Figure 20 for details. None of the individual trials detected a difference in the proportion of patients who still had fever at follow up.

Figure 20: Patients still unwell at follow-up at day 4+ in RDT arm versus clinical diagnosis arm.



Secondary outcomes

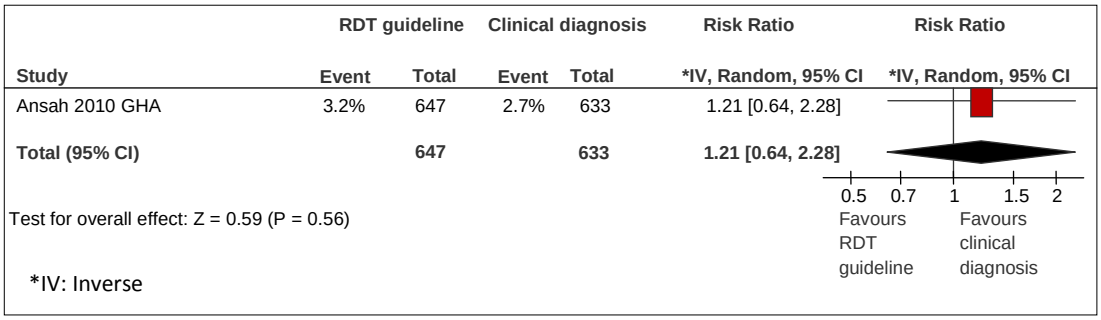
a) Targeting of antimalarials

One study subjected all participants to reference slide microscopy in order to investigate if the use of RDTs in treatment of fever resulted in better targeting of antimalarials. It examined if all malarial patients (slide positive cases) received antimalarials, and whether fewer non-malarial fevers (slide negative cases) received antimalarials after introducing RDTs [36]. The results are summarised below.

Microscopy positive patients not prescribed anti-malarials

No significant difference was detected between RDT and clinical diagnosis arms in the proportion of reference slide positive patients not prescribed antimalarials (RR = 1.21, 95% CI 0.64 to 2.28; one study) [36] (Figure 21 below).

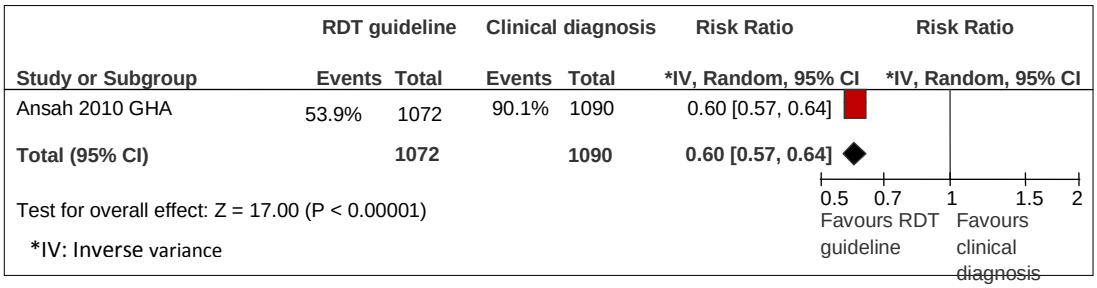
Figure 21: Slide positive cases not prescribed antimalarials in comparison arms



Microscopy negative patients prescribed anti-malarials

Significant difference was detected in the proportion of slide negative cases of fever prescribed antimalarials in favour of RDT-guideline (RR = 0. 60, 95% CI 0.57 to 0.64; one study) [36]. See Figure 22 below.

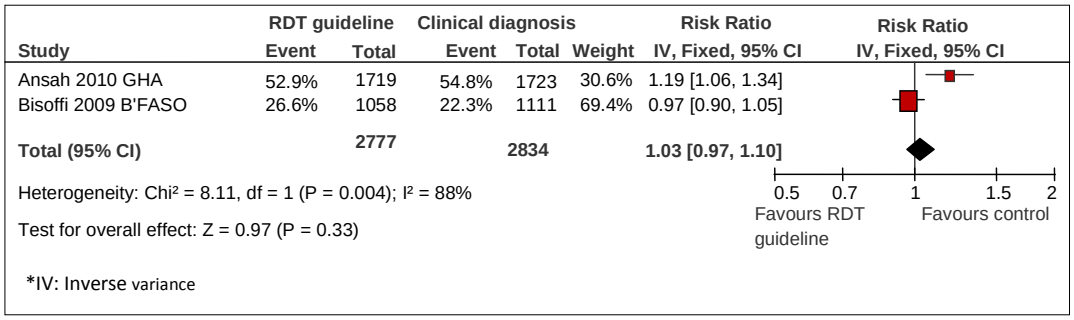
Figure 22: Slide negative cases prescribed antimalarials in the comparison arms



b) Patients prescribed antibiotics

Results from the two trials which evaluated this outcome show considerable variability in the effects of the comparison guidelines on the proportion of participants prescribed antibiotics (I^2 = 88%; p = 0.004) (Figure 23).

Figure 23: Number prescribed antibiotics in the RDT arm versus clinical diagnosis arm



Ansah and colleagues (2010) detected a significant increase in the proportion of participants prescribed antibiotics in the RDT arm [36]. This increase was observed exclusively in older (5+ years) participants in whom prescribing of antibiotics nearly doubled (RR = 1.97, 95% CI 1.69 to 2.30) (not shown in Figure 23). No difference was detected in children (<5 years) (RR = 1.02, 95% CI 0.84 to 1.23). Bisoffi et al detected no significant association between the intervention and the proportion of participants prescribed antibiotics [34].

3.4 Discussion

This chapter reviews evidence from RCTs in order to answer the question whether RDT-based policies can be sufficiently implemented under carefully controlled conditions and whether its application in such conditions can lead to better prescribing and clinical outcomes. Effects are compared relative to treatment policies for fever based on clinical judgement.

3.4.1 Summary of main results

a) Guideline implementation

Overall, HW adherence to guideline was poor. Whereas HWs prescribed antimalarials to virtually all (98% to 100%) participants with positive RDT results, HW response to negative RDT results was generally poor, and varied widely. Adherence to negative RDT results was high in only one trial in which antimalarials were prescribed to only 0.4% of patients with negative RDT results [37]. HW adherence was inadequate in 3 trials, in which as many as 40% to 81% of patients with negative RDTs received antimalarials [34-36]. HW adherence to RDT-based guideline appears to be unpredictable even under relatively optimal conditions.

The reasons for the variability HW adherence between studies are not entirely clear. We know from the literature review that a number of factors can positively influence HW adherence to guidelines, including supervision [93, 112, 117, 146], HW qualifications [90, 112, 146], and training [90, 108, 117, 146]. Further, material or monetary incentives are also known to improve HW performance [153, 154]. Variability in these factors may explain the heterogeneity in HW adherence to RDT test results, which exists between the trials.

In Zambia, on top of training, supply of RDTs and guidelines, the community health workers received a bicycle each and were intensely and regularly monitored [37]. HWs in other studies did not receive any material incentives or similar level of supervision [34-36]. Hence the CHWs in Zambia could have been more motivated to perform than HWs in other trials. Secondly, we know from the literature review that auxiliary staff tends to adhere to guidelines better than professional staff [90, 112, 146]. In Zambia, the attending HWs consisted of community (auxiliary) health workers from health posts [37]. In both Ghana and Burkina Faso, the attending HWs consisted of mostly nurses from dispensaries or health centres [34, 36]; while in Kenya, the HWs consisted of a

variety of cadres of staff from dispensaries, health centres and hospitals [35]. Therefore, variability in settings and qualification of HWs might have partly contributed to the heterogeneity in HW adherence to guidelines between studies.

Although the content of trainings appeared similar in all trials, the quality of delivery could have varied between studies. For example, training was provided for 3 days in each of the trials in Zambia, Ghana and Burkina Faso [34, 36, 37], in Kenya HWs received training for half a day [35]. Thus poor HW performance in Kenya could have been partly due to the short training duration.

Efficacy trials are aimed at optimising performance [40, 98]. On account of the indicators of HW performance analysed, this objective was only achieved in the Zambian trial. The results show that HW performance is likely to vary even under relatively favourable conditions. As a minimum, regular support supervision of HW may be required during the introduction of an RDT-based guideline.

b) Effect on antimalarials prescribed

The evidence indicates that RDT-based policies can reduce the amount of antimalarials prescribed in cases of fever in endemic areas. However, the magnitude of effect of the intervention appears to depend partly on the degree of HW adherence to guidelines, especially on the extent to which they comply with negative test results. The higher the proportion of RDT-negative cases who are prescribed antimalarials in the RDT arm, the smaller is the effect of the intervention on the amount of antimalarials prescribed. Because HW adherence was generally poor, the effect of the intervention on prescribing of antimalarials was also marginal.

Other factors could have also been responsible for the inconsistency in the effect of the intervention. Variability in the prevalence of malaria in the settings where the studies were conducted may have significant influence on the effect of the intervention. The amount of over-use of antimalarials averted by use of RDTs is greater in low malaria prevalence areas than is high prevalence areas [12, 13]. The studies included in this review either did not describe the malaria endemicity of the study contexts [34, 36] or they present results in combined form [35, 37]. Although variability in the age profile of the population is another possible reason for heterogeneity in effect of the intervention

on usage of antimalarials, this review shows that variability in age-group group was not an important factor.

One of the contentious issues about the use of RDTs to guide treatment in fever patients is that it could lead to missed diagnosis of malarial cases [26, 70]. The review detected no difference between RDT and clinical diagnosis arms in the proportion of reference microscopy positive patients missing anti-malarials. However, this outcome was examined in one trial only [36]. The same trial shows that use of RDTs can significantly avert inappropriate prescribing of antimalarials [36].

c) Effect on prescribing of antibiotics

The review reveals considerable variability in the effects of the comparison guidelines on the proportion of participants prescribed antibiotics. While the trial in Ghana detected a significant increase in the proportion of participants prescribed antibiotics in the RDT arm [36], the one from Burkina Faso detected no difference between the comparison arms [34]. Overall, effect of the intervention under controlled conditions is still unclear.

Usage of antibiotics is thought to be an indicator of how parasite negative patients are managed [138, 155]. Variability in prescribing of antibiotics might reflect variability in the way RDT-negative patients are managed in different settings. Generally, users of RDT-based guidelines appear to be unclear about how to manage parasite-negative patients [32, 34-37, 120, 121, 151]. However, this finding might also reflect variability in availability (i.e. shortage) of essential antibiotics.

d) Effect of clinical outcomes

The review did not demonstrate a difference between RDT and clinical diagnosis arms in the proportion of participants still symptomatic at follow-up evaluation (4 - 7 days after treatment).

This could be due to a variety of factors. Firstly, differences in outcomes may be small. In order to detect small differences in effects, randomised trials require large sample sizes. For the studies included in this review, sample size calculations were based on anticipated differences between groups in prescribing of antimalarials—which may be bigger than differences in clinical outcomes. Therefore, sample sizes used in these studies may have been too small to detect differences in clinical outcomes. Secondly, in Bisoffi et al. (2009), clinician adherence to RDT-based guidelines was poor [34].

Therefore, the difference in the quality of prescribing provided in the two arms may not have been significantly different.

Thirdly, we know from the literature review that outcomes data are usually less sensitive in measuring quality of care [85]. This is especially true if measures are implicit, such as the subjective measures used in this study. Differences in outcomes may not be detected even if there is improvement in the quality of the process of care. On the other hand, differences in outcomes may be due to antecedent factors [85].

3.4.2 Comparison with other studies or reviews

Non-randomised trials have also shown that use of RDT-based guidelines in routine practice can reduce prescribing of antimalarials in settings where presumptive treatment is the norm, especially in low transmission areas [32, 120, 121]. A weekly cross-over trial from Zanzibar (Tanzania) detected a reduction (from about 5% to 2.5%) in the risk of persistent symptoms after introducing RDT guidelines (OR 0.5, 95% CI 0.3 to 0.9, P = 0.005)[121]. These studies did not evaluate the effects of RDT-based policies on prescribing of antibiotics. Further randomised controlled trials are required to enable a robust conclusion about the effects of RDT-based policy on prescribing of antibiotics and clinical outcomes.

3.4.3 Limitations

a) Use of prescribing as a measure of quality

In this chapter, analysis of HW adherence is based on prescribing of anti-malarials relative to RDT results. RDT-based guidelines are aimed at improving the quality of the process of care as a whole, including clinical assessment, RDT performance and patient counselling. Notably, the quality of clinical assessment is crucial in the differential diagnosis of the causes of RDT-negative fevers, and in targeting antibiotics [19, 70-72]. Use of prescribing data alone as a measure of adherence to guidelines fails to fully describe the overall quality of care in fever patients, and the quality of the diagnostic processes of care specifically [109, 113, 149]. Additionally, prescribing data are known to over-estimate the overall quality of care [62, 90-92, 108-113, 115-117]. Therefore, HW performance may have been much poorer in the studies analysed than is portrayed by the data on prescribing of antimalarials.

b) Heterogeneity

The trials analysed vary considerably in design, number of supplementary interventions, implementation, and settings. This probably reflects the limitations of using archetypical RCT design to attribute effects to a complex public health intervention. Variability in methods and settings hampers the ability to compare results from these trials.

c) Design effect

Data on prescribing tend to be highly correlated because they reflect the similarity in the quality of care offered to several cases by a few HWs. Randomisation, whether at individual or cluster level, may not address clustering of prescribing data, especially when the number of facilities enrolled is small, and there are a few HWs per facility [102]. When analysis is undertaken at patient level, clustering (design effect) reduces the power of the study to detect differences in effect between groups, and results in over-precise estimates—except if appropriate adjustments are made [99, 102]. Of the studies included in this analysis, cluster randomised trials took into account potential clustering of effects in the calculation of sample size—thereby recruiting 30 and 31 health facilities respectively. Individual randomised trials did not take clustering into account, and recruited 3 and 10 health facilities respectively. Therefore, in the latter group of trials, effects of the intervention on prescribing of antimalarials and antibiotics may have been masked by clustering.

d) Use of RCTs in a complex intervention

The primary purpose of an RCT is to attribute outcomes to interventions and to explain how interventions work. RCTs are generally unsuitable for attributing effects to elements of a complex intervention because it is difficult to identify the components of the intervention which is responsible for an observed effect [74]. This review detected variability in implementation of guidelines, which potentially could help explain differences in outcomes. However, the trials lacked data on most of the process attributes that could have been used to demonstrate causal relationship. Therefore, while we know how HW adherence to RDT results impact on usage of antimalarials, it is unclear how implementation of other components of the guideline could impact on other outcomes indicators—notably prescribing of antibiotics and clinical outcomes.

e) Strict selection criteria

The research quality criteria used to select studies for inclusion in this review may have been overly biased against studies with large effects of the intervention, but whose designs were deemed to be sub-optimal as per the selection criteria, for example D'Acremont et al (2011), Kyabayinze et al. (2011) and Msellem et al. (2009) [120, 121, 156]. On the hand, the criteria may have been biased towards interventions with marginal effects because the designs of the studies met the selection criteria, for example Skarbisnki et al (2009) [35]. Consequently the studies reviewed are too few to support a robust conclusion about the effects of RDT-based guidelines relative to clinical diagnosis. Further, the combined intervention effects may be smaller than would have been if quasi-experimental studies were also included in the review.

f) External validity

The trials included in this review were conducted under relatively favourable conditions characterised by standardised interventions [34, 36, 37], well resourced settings [34-37], motivated HWs [37], homogeneous population [35, 37], and intense monitoring of intervention [37]. As such the evidence from the individual trials may not be generalisable to the general population in routine practice [98]. In the context of this thesis, knowledge of efficacy is necessary to aid interpretation of evidence from the effectiveness trial in chapter 4 [40]. The external validity of the evidence in this chapter is not a primary objective of this thesis. Yet generalisable conclusions can still be drawn from this chapter because the systematic review pools together similar effects in a variety of populations [98]

g) Use of outcomes indicators to measure quality

Outcomes data are usually a less sensitive measure of quality of care. This is especially true if measures are implicit, such as the subjective measures used in this study. Differences in outcomes may not be related to the factors under the control of the HW. Rather, they might be due to the influence of antecedent factors [85]. Therefore, a lack of improvement in clinical outcomes should not be considered to indicate lack of improvement in the overall quality of care [85].

3.5 Authors' conclusions

Policies based on RDTs instead of clinical diagnosis in fever management may significantly reduce usage of antimalarials if applied under optimal conditions. However, HWs may not adhere to the guideline sufficiently enough to result in significant improvements in prescribing of antimalarials, even under these conditions which are designed to optimise performance. We are also uncertain about the effect of the policy on prescribing of antibiotics and on clinical outcomes. Therefore, if a lack of effect is detected in the effectiveness (pragmatic) trial in chapter 4, we can not know if it is because the intervention is, in fact, inefficacious, or whether it is due to inappropriate design or invalid measure; or if it is due to poor implementation.

3.5.1 Implications for research

More efficacy trials may be required in order to make a robust conclusion about the effects of the intervention on clinical outcomes. Trials need to be more standardised in design, interventions, implementation and settings so that results are more comparable. Trials need to take design effects into account, even for prescribing outcomes. Perhaps more explicit indicators are needed to measure clinical outcomes.

Studies evaluating effects of the intervention on clinical outcomes need to include an assessment of implementation of all components of the guideline. Given the limitations of RCTs in assessing complex public health interventions, quasi-experimental can be more useful for attributing effects to a multi-component intervention. When incorporated into a model, variation in implementation of various components of a guideline may be used to predict specific outcomes attribute [85]Future studies should enrol larger numbers of health facilities, and should analyse results by malaria endemicity in the study settings to facilitate comparison of results.

Chapter 4

Analysis of data from a pragmatic trial

Chapter 4 Effect of RDT-based policy on quality of care in routine practice

4.1 Introduction

In this chapter I present the analysis of an existing dataset from a pragmatic trial in Uganda in order to determine if application of an RDT-based guideline through routine practice can lead to improved prescribing, improved clinical outcomes and lower healthcare cost relative to presumptive treatment. Additionally, it describes the quality of implementation of the RDT-based guideline for Uganda in routine practice—in terms of HW's use of RDTs, diagnostic accuracy of RDTs and HW response to RDT results. Further, it evaluates the impact of variability in malaria endemicity and age profile on the magnitude and direction of effects of the new policy. The evidence from this chapter can inform judgement as to whether it is worthwhile, from both economic and clinical perspectives, to scale-up the use of RDT-based policy for fever in all age groups in all malaria transmission settings in Uganda; or whether to target its use to specific age groups and transmission settings.

The dataset analysed in this chapter comes from a trial which was designed to evaluate the diagnostic accuracy of RDTs, and to compare the effectiveness of an RDT-supported guideline for fever versus presumptive treatment in government HCs in Uganda. It was carried out in 2008 in high, medium and low malaria transmission settings. The trials was conducted by the Uganda Malaria Surveillance Project (UMSP)—a collaborative initiative comprising researchers at Makerere University in Uganda, the University of California and San Francisco in the US, and the Uganda Ministry of Health (MOH) (see protocol, Appendix 9).

This dataset had, up to the time of this analysis, not been analysed and results presented elsewhere.

4.1.1 Research question

Does the evidence from a pragmatic trial in Uganda show improved quality of patient care when treatment policy for fever based on RDT results is applied to a large population through routine health services?

4.1.2 Aim

To review evidence from a pragmatic trial to determine if RDT-based policies for fever lead to better quality of patient care than presumptive treatment when applied in routine practice

4.1.3 Objectives

To use the existing data set to:

1. Describe the quality of guideline implementation in routine practice
2. Compare effects of a treatment policy for fever based on RDTs versus a policy based on clinical diagnosis in relation to:
    a. prescribing of anti-malarials and antibiotics
    b. clinical outcomes
    c. cost
3. Assess if a treatment policy for fever based on RDTs leads to better targeting of antimalarials than a policy based on clinical diagnosis.

The analysis takes place in settings with high, medium and low malaria transmission intensity. Where relevant, results are analysed for different age groups (< 5 years versus 5+ years).

4.2 Methods

This section briefly describes the main features of the UMSP trial, and outlines my research strategy for analysing the dataset. A more detailed description of the trial which generated the database is found in Appendix 9.

4.2.1 Trial design

a) Setting

The trial was conducted in 3 districts representing 3 malaria transmission settings: Tororo (High transmission), Jinja—peri-urban setting (medium transmission) and Mubende (low transmission setting). Table 18 below offers further descriptive information about the study sites.

Table 18: Characteristics of the Study Sites

	DISTRICTS		
	TORORO	JINJA	MUBENDE
Location	Rural, E. Uganda	Peri-urban, S.E Uganda, near Lake Victoria	Rural, S.W. Uganda
Malaria transmission intensity	High	Medium	Low
Landscape	Plane, dry savannah grassland	Hilly grassland	Hilly grassland

Sources: [75]

b) Design and Data Inventory

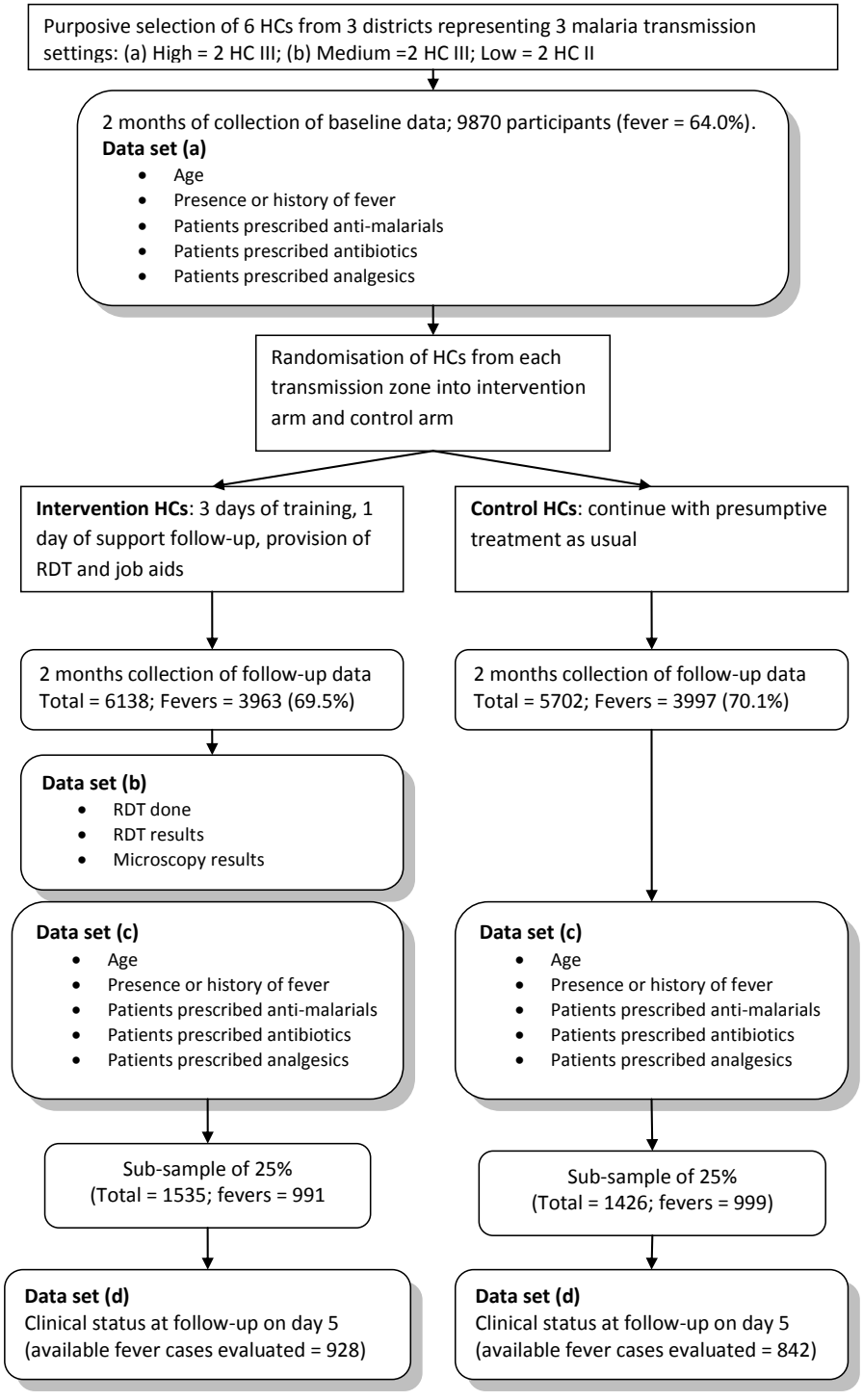
Figure 24 below depicts the trial design and the sets of data used in the analysis.

Health facilities

In each of the 3 districts, the team identified two health centres (HCs) for the evaluation, all of which without the capacity for microscopy. In both Tororo (high transmission) and Jinja (low transmission), the pairs of HCs identified had referral level (HC IIIs), while the pair enrolled in Mubende (low transmission) were HC IIs (referral level II status). The recommended staffing norm for Uganda is shown in Appendix 2. A typical HC III is manned by at least 3 staff, comprising any of the following cadres: clinical officers, registered nurses/midwives, enrolled nurses/midwives, nursing assistants and laboratory technician/assistants. A typical HC II is manned by at leat 2 staff consisting of any of a nurse or midwife or nursing assistants. According to the protocol (see Appendix 9) each of the HC selected (a) lacked basic microscopy services, (b) had at least 3 clinical staff, (c) had a workload of at least 500 patients per month and (d) was within radius of Uganda Malaria Surveillance Project (UMSP) sentinel sites.

Treatment of malaria was based on clinical judgement at all sites prior to the study. The HCs were randomly assigned to the intervention and control arms. At each HC, data were collected in 2 phases: before and after the intervention, with each phase lasting 2 months.

Figure 24: Trial design and data inventory



Participants

The trial recruited all ambulatory patients presenting at the facilities, except those with severe conditions, or those who did not consent. However, for the purposes of this report, analysis of effects of the intervention was based on participants presenting with fever, with or without other symptoms.

Baseline survey

The trial was preceded by 2 months of baseline survey during which surveillance data were collected on every outpatient with regard to their age, gender, symptoms, diagnosis, and treatment prescribed. Follow-up evaluations of patients' health outcomes were NOT conducted during the baseline survey. In addition, there was no parasitological confirmation and typing of malaria during the baseline survey.

Interventions

Following the baseline survey, one of the health centres in each district was randomly selected to receive to receive the intervention. The intervention consisted of: (a) 3-day training in the use of RDTs and RDT-based guideline, (b) provision of RDT kits (*Paracheck* HRP2-based test, manufactured by Orchid Biomedicals (Goa, India), (c) provision of guidelines, and (d) additional support training, 2 weeks after the initial training. No additional supplies were provided apart from RDT kits and guidelines. Once data collection commenced, no additional support training and supervision was provided as part of the trial. Health workers selected the patients for RDTs, performed the RDTs and read the results of the test. There was no interference with their decisions on who to test, and how to respond to the test results.

The control HCs in each district did not receive similar training, or RDT kits and related job aids, and continued with presumptive treatment as before.

Post-intervention survey

Surveillance data collection (age, symptoms, RDT performance, diagnosis, prescriptions) continued for 2 more months post-intervention. A sub-sample of the study participants (25%) were selected through systematic random sampling from each intervention and control HC. They were followed-up and their clinical outcomes evaluated at day 5 post treatment initiation. Outcomes were based on self report, recorded as one of the following: (a) Improved; (b) No change; (c) Worse; (d) Unable to state

Quality control

At the time of RDT preparation, a research assistant prepared blood smear for reference microscopy. The research assistant stained the smears with 2% Giemsa for 30 minutes, mounted them with DPX (distyrene, plasticizer, xylene mountant) and cover slips, and stored them in slide boxes. The smears were read at a UMSP reference laboratory in Kampala by experienced technicians, who were blinded to the results of the patients' RDT results. Smears were evaluated for the presence of parasitaemia (asexual forms) and gametocytes. A slide was considered negative when examination of 100 high power fields did not reveal asexual parasites or gametocytes. For quality control, all slides were re-read by a second microscopist, and a third reviewer settled any discrepant readings.

4.2.2 Summary of data collected

Table 19 below summarises the data used in the analyses that follow

Table 19: Data used in analysis

Data	Study arms	Source (survey)	Objectives in which data are used
• RDT done • RDT results • Microscopy results	Intervention HCs only	Follow-up *(dataset (b))*	Objective 1 Objective 3
• Age • Presence or history of fever • Anti-malarials prescribed • Antibiotics prescribed • Analgesics prescribed	Intervention and control HCs	Baseline and follow-up *(datasets (a)+(c))*	Objective 2 Objective 3
• Clinical status on day 5	Intervention and control HCs	Follow-up *(data sets (d))*	Objective 2

4.2.3 Analyses undertaken in this chapter

This section shows how the variables in objectives 1 to 3 were measured using datasets (a), (b), (c) and (d). The analyses focus on participants with fever only.

Table 20: Summary of the analyses carried out in chapter 4

Variables	Analyses
Objective 1 (implementation)	Use of RDTs and prescribing of antimalarials relative to RDT results Diagnostic accuracy of RDTs (sensitivity and specificity)
Objective 2 (primary outcomes)	Change in prescribing of antimalarials and antibiotics after intervention Change in cost after intervention Difference in clinical outcomes between intervention and control HCs
Objective 3 (secondary outcomes)	Slide positive cases missed by RDTs Slide negative cases prescribed antimalarials

4.2.4 Quality of guideline implementation

The thesis uses data sets describing the use of RDTs and microscopy (dataset (b)) at intervention HCs to assess two critical aspects of implementation of RDT-based policies, namely HW performance (adherence to guideline) and diagnostic accuracy of RDTs. Because of data limitation, this analysis does not evaluate the quality of implementation of other components of the guideline such as medical history, clinical examination and advice on medications. Implementation of all components of the guideline is assessed in Chapter 5 using a Health Facility Assessment (HFA) tool.

a) HW performance

Because of data limitation, judgement of HW performance is based on assessment of 3 tasks only: (a) use of RDTs, (b) prescribing of antimalarials in response to positive RDT results, and (c) prescribing of antimalarials in response to negative RDT results. HWs were expected to perform each of the 3 tasks correctly at least 95% of the time. Otherwise implementation of the guideline or HW performance was judged to be inadequate [104, 106, 127, 128, 140]. The 95% threshold has been used in several health service surveys [104, 106, 127, 128]. An economic model developed by Lubell and colleagues in 2008 also suggests that the economic advantage (cost-saving) gained by using HRP-2-based RDTs over presumptive treatment is lost once HW compliance with RDT-based guidelines falls below 95% in children (<5 years) and/or in high prevalence areas [140].

b) Diagnostic accuracy of RDT test

Sensitivity, specificity, false positive and false negative error rates are calculated by cross-tabulating RDT results with microscopy results. Sensitivity and specificity of RDT at a given site was considered high if the respective measure had a value of at least 95% and 90% [20, 22, 139].

4.2.5 Primary outcomes

a) Change in prescribing of anti-malarials and antibiotics

This objective assesses the effect of implementing an RDT-based guideline on prescribing of anti-malarials and antibiotics—the 2 most commonly prescribed items in this trial, the 2 items with the greatest implications for cost [12, 13, 33, 140] and the 2 items at the centre of debate [26, 138].

Effect is measured as an absolute change from baseline (within 95% CI). It is calculated by subtracting the percent of participants prescribed the relevant drugs after introducing RDTs (dataset (c)) from the percent at baseline (data set (a)). At each intervention setting, change in prescribing is calculated for the pair of intervention and control HCs. The results at an intervention HC in a given transmission setting can be compared with that at a corresponding control HC to rule out the effects of extraneous factors on the results observed at the intervention HC[5].

b) Calculating change in costs

Perspective: Cost was calculated from the perspective of the provider of care.

Inputs: I included the cost of biomedical consumables that were used frequently in the diagnosis and treatment of fever cases (data sets (a) and (b)); namely anti-malarials, antibiotics and antipyretics or analgesics.

Table 21 below shows the items analysed, the corresponding dosage requirements, unit prices and the cost for a full dose required for an average individual in a particular age category.

---

[5] To rule out the role of extraneous factors occurring between pre-test and post-test, which can affect the outcome under study, e.g. (a) history: any external event such as stockout of anti-malarials (b) maturation: internal changes such as exposure to knowledge (122. Bonate, L.P., *Aanalysis of pretest-posttest designs*2000: Chapman and Hall/CRC.

Table 21 Recommended quantities, unit prices and cost of included inputs

Items	Required quantity per case (average)	Unit price* US$ (2008)	Cost per case** US$ (2008)
RDT kits			
<5	1	0.6	0.6
5+	1	0.6	0.6
All	1	0.6	0.6
ACT			
<5	6	0.084	0.504
5+	24	0.084	2.016
All	24	0.084	2.016
Antibiotics: septrin			
<5	3	0.01	0.03
5+	20	0.01	0.2
All	20	0.01	0.2
Antibiotics: others			
<5	6	0.02	0.12
5+	40	0.02	0.8
All	40	0.02	0.8
Weighted cost of antibiotics***			
<5			0.104
5+			0.692
All			0.692
Analgesics			
<5	3	0.003	0.009
5+	18	0.003	0.054
All	18	0.003	0.054
Source:	STG, MOH	NMS	

* per tablet/capsule or test
** assuming full dose as per guideline
***septrin tablets were prescribed 1.6 times more often than other (6-hourly) antibiotics. A weighted average cost of antibiotics are used in analysis of cost, taking into consideration the relative frequency of prescribing of the two categories of antibiotics

STG: Standard Treatment Guideline
MOH: Ministry of health user manual (RDT-based guideline), Uganda
NMS: National Medical Stores, Uganda

In this table, some of the items (e.g. analgesics, "other antibiotics") are grouped together. The list of the full range of specific items[6] included in the cost analysis, the corresponding dosage requirements and unit prices are shown in Appendix 10 and Appendix 11.

Items that were occasionally prescribed, or were not typically used in the diagnosis and treatment of fever cases were excluded (e.g. vaccines, Vitamin A supplements, anti-helminths). Other programmatic costs (e.g. staff time, the cost of training and support supervision) were also excluded because they were not contained in the database, and the responsible key informants were not available to provide the necessary information.

Calculating cost: The cost of any item prescribed to an individual was calculated as:

$$Cost = quantity \times unit\ price$$

Data sets (a) and (c) only show the types of biomedical consumables prescribed for each study participant. They do not show the quantity prescribed or whether the participant received all the required items. The analysis assumed that all the identified items were prescribed according to the standard treatment guidelines [157, 158], and that all the items were received.

Since the regimens for medicines are age dependent, quantities are computed by age groups. For each item prescribed to an individual in a particular age group, I assumed the dosage requirement for the average age of an individual in that age group, as shown below

Table 22: Average age of participants by age-groups

Age group	Average age
a) < 5 years	1.6 years
b) 5+ years	25 years
c) All	18.5 years

The total cost (for the recommended quantity) of an item that was prescribed to a participant in a given age group is obtained by multiplying the required quantity for that individual, by the unit price of the item. Prices were extracted from the National

[6] E.g. antimalarials: ACT, quinine, SP

Medical Stores (NMS) catalogue and price indicator of 2008—which represents the official government prices of the items for the period under review [159]. Prices were captured in Uganda Shillings (UGX), and subsequently converted to 2008 US$, at an exchange rate of UGX 2177.56—the average exchange rate in 2008 as per Bank of Uganda [160].

Measure of effect: Effect is measured as absolute change from baseline (within 95% CI). This is calculated for every study site by subtracting the post-intervention cost from the baseline cost. The results show the amount of money added to, or saved by the healthcare system, by using RDT-supported guideline to treat a case of fever instead of presumptive treatment. Change in cost is calculated for both intervention and control HCs. This allows results from intervention HCs to be compared across 3 transmission settings to assess if they are significantly different. Further, the result from an intervention HC in a given transmission setting can be compared with that at a corresponding control HC to rule out the effects of extraneous factors on the results observed at the intervention HC.

c) Difference in clinical outcomes between groups

Data on clinical outcomes is provided for 25% of participants. Measurement of outcomes was based on self-reports, and were recorded as (a) Improved; (b) No change; (c) Worse; (d) Unable to state. The frequencies for health states (b), (c) and (d) were low at most HCs. Therefore, for the purpose of this analysis, responses are grouped into 2 categories: "improved" and "not improved", where the latter comprises health states (b), (c) and (d).

Assessment of clinical outcomes was carried out at both intervention and control HCs during the follow-up surveys. Since clinical outcomes were not assessed at baseline, this analysis does not make a before-and-after comparison as is done in the analysis of change in prescribing and cost. Instead, it compares clinical outcomes between each pair of intervention and control sites using data set (d). For each pair of HC, I computed the percent of patients reporting improvements in their conditions 5 days after initiation of treatment. The effect of using a RDT-supported guideline at each transmission zone is measured as the ratio of the percent of patients reporting improvement in the intervention arm to the percent in the control arm (i.e. RR, 95% CI).

4.2.6 Secondary outcomes

a) Missed cases of malaria after introducing RDTs

Sub-analysis was undertaken to determine the percent of reference slide positive cases who were not prescribed antimalarials at the intervention HCs after introducing RDTs. The analysis is undertaken by cross-tabulating the variables "any antimalarials" versus "microscopy results". The percent of slide positive cases who were not prescribed any antimalarials is compared with the false negative (β) error rates at the same HCs. The results show the impact of diagnostic inaccuracy of RDTs on the number of malarial cases not prescribed any antimalarials.

b) Slide-negative fevers given antimalarials after introducing RDTs

The analysis described in (a) above also generates the percent of reference slide-negative cases prescribed antimalarials at the intervention HCs. These results are compared with the false positive (α) error rates at same HCs. They show the impact of diagnostic inaccuracy of RDTs on the proportion of slide negative (non-malarial) fevers that are still prescribed antimalarials despite use of RDT test.

4.2.7 Statistical analysis

Analysis of primary outcomes was carried out using SPSS (PASW statistics 18) to compute summary statistics (percent, means and medians) and Revman 5.1 to generate confidence intervals (CIs) around measures of effect (differences or risk ratios). For computing 95% CIs around summary statistics (e.g. percent, averages)—such as in baseline characteristics, clinician adherence, accuracy of RDTs— analyses were carried out using MS Excel 2007 as follows.

CIs around proportions were calculated using the formula for CI for single proportions [147], namely:

$$\text{CI} = p \pm z^*\sqrt{pq/n}$$

Where: z = z-score (1.96 for 95% CI)

p = event rate (percent)

q = (1-p)

n = sample size

CIs around arithmetic means for a sample population were calculated by substituting the values of standard error (SE) in the following formula.

CI = mean ± z*SE;

- where z = z-score (1.96 for 95% CI)

Means and SEs were generated using SPSS (PASW statistics 18)

4.3 Results

4.3.1 Characteristics of the study population

The UMSP trial was located in 3 districts representing 3 malaria transmission settings: Tororo (high), Jinja, peri-urban setting (medium), and Mubende (low). Six health centres (HCs), 2 from each district, were randomised to the intervention or control arm. The comparison HCs in each district were similar in characteristics: each district pair was of level II or III referral status; each HC had at least 3 clinical staff, and each had a workload of ≥ 500 patients per month. None of the HCs was capable of performing blood slide microscopy for malaria. Table 23 below summarises the study participants.

Table 23: Study participants

Variable*	District	Setting	Intervention HCs		Control HCs	
Baseline survey						
Total number of participants (N)	Tororo	High	1653		2130	
	Jinja	Medium	1843		1626	
	Mubende	Low	1394		1224	
Age of participants with fever (yrs): median (Q1, Q3)	Tororo	High	14	(2, 27)	12	(2, 27)
	Jinja	Medium	19	(9, 30)	18	(7, 29)
	Mubende	Low	20	(12, 30)	20	(11, 32)
Participants with fever: N(%)	Tororo	High	985	(59.6%)	1446	(67.9%)
	Jinja	Medium	1006	(54.6%)	1168	(71.8%)
	Mubende	Low	942	(67.6%)	771	(63.0%)
Follow-up survey						
Total number of participants (N)	Tororo	High	3209		2668	
	Jinja	Medium	1915		1480	
	Mubende	Low	1014		1554	
Age of participants with fever (yrs): median (Q1, Q3)	Tororo	High	5	(1.4, 22)	10	(1.7, 25)
	Jinja	Medium	15	(4, 28)	15	(4, 29)
	Mubende	Low	20	(10, 30)	21	(8, 28)
Participants with fever: N(%)	Tororo	High	2288	(71.3%)	1909	(71.6%)
	Jinja	Medium	1073	(56.0%)	1140	(77.0%)
	Mubende	Low	602	(59.4%)	948	(61.0%)
Fever patients with positive blood slides for malaria: % (95% CI)	Tororo	High	53.5	(53.2, 57.4)	-	
	Jinja	Medium	36.5	(33.5, 39.5)	-	
	Mubende	Low	29.3	(25.6, 33.0)	-	

*N =total number; Q1, Q3 = 1st and 3rd quartiles; CI = confidence interval

A total of 21,710 participants were enrolled during the 4-month study period (2 months of baseline survey and 2 months of post-intervention data collection): 9870 during baseline and 11,840 during the follow-up surveys. There was no systematic difference between the study arms in change (increase or decrease) in enrolment during the follow-up survey (see Table 23).

During both surveys, the participants enrolled in Tororo (high transmission) were generally younger than those enrolled in Jinja (medium transmission) and Mubende (low transmission). There was no difference in participants' age between any pair of comparison HCs. Additionally, at each study site, there was no difference in the age of participants enrolled at baseline versus those enrolled during the follow-up survey; the notable exception was the intervention site in Tororo (high transmission), where the participants recruited after the intervention were much younger than those recruited at baseline (5 years vs. 10 years respectively).

More than 65.8% of the participants (14283) had fever, with notable variability in the prevalence of fever between sites and within sites (before vs. after intervention).

The prevalence of malaria in fever was assessed at intervention sites only, and only during the follow-up survey[7]. Overall, the percent of fever patients with positive blood slides at the intervention centres was 53.3% in Tororo (high transmission), 36.5% in Jinja (medium transmission) and 29.3% in Mubende (low transmission). The prevalence of malaria in fever patients was higher in younger participants than in older counter parts, being 2 times higher in a high transmission setting, and 1.6 times higher in medium and low transmission settings (not shown in Table 23).

4.3.2 Baseline values of study variables

Baseline data were not analysed prior to randomisation, and were not used in matching the HCs. In this section I present the baseline values for 3 outcome variables at the various study sites: (a) patients prescribed any anti-malarials, (b) patients prescribed antibiotics and (c) the average cost of diagnosing and treating a case of fever. The aim is to show any similarities and/or imbalance, at baseline, between comparison HCs. Data

[7] Reference microscopy was performed at the intervention sites after introducing RDT-based guideline, to assess the accuracy of RDTs

on clinical outcomes are not presented here because the relevant data were not captured during baseline.

Table 24: Baseline values of study variables

			Control HCs		Intervention HCs	
Variable	**District**	**Setting**	Value	(95% CI)	Value	(95% CI)
Patients prescribed any anti-malarials (%)	Tororo	High	88.7	(87.1, 90.4)	92.9	(91.3, 94.5)
	Jinja	Medium	97.3	(96.4, 98.4)	93.3	(91.8, 94.9)
	Mubende	Low	99.5	(99.0, 100.0)	99.4	(98.9, 99.9)
Patients prescribed any antibiotics (%)	Tororo	High	53.1	(50.5, 55.7)	43.7	(40.6, 46.8)
	Jinja	Medium	55.1	(52.2, 58.0)	50.7	(47.6, 53.8)
	Mubende	Low	39.5	(36.0, 43.0)	51.4	(48.2, 54.6)
Cost of diagnosing and treating 1 case of fever(2008 US$)	Tororo	High	1.44	(1.39, 1.49)	1.47	(1.41, 1.53)
	Jinja	Medium	1.97	(1.92, 2.01)	1.86	(1.81, 1.92)
	Mubende	Low	2.07	(2.02, 2.13)	2.04	(1.99, 2.09)

Overall 94.5% of fevers were prescribed anti-malarials at baseline. The percent of cases prescribed anti-malarials was similar across the 3 transmission zones. In Mubende (low), virtually all participants were prescribed antimalarials at baseline. The percent of cases prescribed antimalarials in Tororo (high transmission) was significantly lower. Statistically significant differences were detected between the pair of HCs in Tororo (high) and Jinja (medium) in the amount of anti-malarials prescribing at baseline. No significant imbalance was detected between the pair in Mubende.

About half (49.7%) of fever cases were prescribed any antibiotics at baseline. There were significant imbalances between the comparison HCs in Tororo (high) and Mubende (low) in the percent of cases prescribed antibiotics. No significant imbalance was detected between the pair in Jinja (medium). No systematic pattern in prescribing of antibiotics was detected relative to malaria transmission settings.

The cost of diagnosing and treating a case of fever at all the study sites ranged from just under US$ 1.45 to just over US$ 2.0, with an overall average of US$ 1.78 (2008US$). The lower the malaria transmission intensity, the higher was the average cost. There was no baseline imbalance in cost between any pair of comparison HCs.

4.3.3 Quality of implementation of RDT-based guideline

This section describes the quality of implementation of RDT-supported guideline in terms of 2 critical success factors:

(a) Clinician adherence with the requirement to test all suspected cases of malaria (fever) prior to treatment; and to restrict anti-malarials to those with positive RDT tests only
(b) Accuracy of the RDT results (sensitivity, specificity, α- and β-errors).

Clinician adherence

Table 25 below summarises HWs use of RDTs and response to the results in patients presenting with fever.

Table 25: Use of RDTs, HW response to RDT results, and diagnostic accuracy of RDTs

			Intervention HCs		
Variable	**District**	**Setting**	N*	%	(95% CI)
Patients tested with RDT	Tororo	High	2221	97.3	(96.6, 98.0)
	Jinja	Medium	1044	97.9	(97.0, 98.8)
	Mubende	Low	594	98.7	(97.8, 99.6)
Patients with positive RDT & prescribed anti-malarials	Tororo	High	1617	99.4	(99.0, 99.8)
	Jinja	Medium	475	98.7	(97.7, 99.7)
	Mubende	Low	189	100	(100, 100)
Patients with negative RDT & prescribed anti-malarials	Tororo	High	604	0.7	(0.0, 1.4)
	Jinja	Medium	569	0.0	(0.0, 0.0)
	Mubende	Low	405	0.5	(0.0, 1.2)

**N: number of febrile patients analysed*

At least 97% of all fever cases at each intervention site were tested with RDTs prior to treatment. In addition, virtually all (98.7% to 100%) fever cases that tested positive on RDT were prescribed anti-malarials. On the other hand, practically none (<0.8%) of those that tested negative on RDT received any anti-malarials. These results show that HWs complied adequately with all the requirements of the guidelines regarding use of RDTs and prescribing of antimalarials. Accordingly, the clinicians are judged as having shown high adherence to these components of the guideline.

Diagnostic accuracy of RDTs

Table 26 displays the performance of RDTs against reference microscopy in febrile patients, by transmission settings.

Table 26: RDT results versus reference microscopy results in patients with fever

			Results of reference microscopy	
			+	-
Tororo (High)	RDT results	+	98.0% (n =1163)	43.9% (n = 451)
		-	2.0% (n = 24)	56.1% (n = 577)
		Total	1187	1028
Jinja (Medium)	RDT results	+	95.5% (n = 341)	15.2% (n = 94)
		-	4.5% (n = 16)	84.8% (n = 523)
		Total	357	617
Mubende (Low)	RDT results	+	82.8% (n = 144)	10.5% (n = 144)
		-	17.2% (n = 30)	89.5% (n = 374)
		Total	174	418

n = number of events; % refers to the proportions of cases detected by RDT out of the column total

The sensitivity of RDTs was high in high transmission (98.0%; 95% CI: 97.4% to 98.6%) and medium transmission (95.5%; 95% CI: 94.3% to 96.7%) transmission settings; it was low (82.8%; 95% CI: 79.8% to 85.8%) in low transmission setting. Accordingly, false-negative error rates of RDT were 2%, 4.5% and 17.2% at high, medium and low transmission sites respectively. Specificity of RDT varied between high and low (89.5%; 95% CI: 87.5% to 91.0%) in low transmission setting. The specificity of RDTs was low in both medium (84.8%) and high (56.1%) transmission settings. Accordingly, false positive error rates of RDT were 10.5%, 15.2% and 43.1% in low, medium and high transmission settings respectively.

Therefore, implementation of the RDT-guideline was characterized by high clinician adherence to the recommendations of the guideline. It was also characterised by a high false positive (α-error) rate at the high transmission site, and by high false negative (β-error) rate at the low transmission site.

4.3.4 Change in prescribing of antimalarials after intervention

Table 27 below displays the amount of change in the proportion of fever patients prescribed anti-malarials at intervention sites after introducing RDT-based guidelines. Additionally, it displays similar results for corresponding control sites. The summary data showing the number of patients prescribed antimalarials before and after the intervention are shown in Appendix 12.

Table 27: Change in proportion of patients prescribed antimalarials after intervention

		Intervention HCs			Control HCs		
District	**Setting**	N*	%	(95% CI)	N*	%	(95% CI)
Tororo	High	3273	-22.1	(-24.6, -19.6)	3355	5.9	(4.0, 7.8)
Jinja	Medium	2079	-48.9	(-52.3, -45.5)	2308	1.0	(-0.2, 2.2)
Mubende	Low	1544	-60.2	(-64.1, -56.3)	1719	-0.2	(-0.9, 10.5)

*N: total number of febrile participants analysed at each site, before and after intervention. Summary statistics showing number analysed and number prescribed antimalarials before and after intervention is shown in Appendix 12

The analysis shows that prescribing of anti-malarials reduced significantly at all the intervention sites after introducing RDT-based guidelines. In addition, it shows significant differences between the transmission zones in the decline in the usage of antimalarials. The percent of patients prescribed anti-malarials declined by 22.1% in Tororo (high transmission); by 48.9% in Jinja (medium transmission) and by 60.2% in Mubende (low transmission). Therefore, the lower the malaria transmission intensity, the higher was the decline in the number of patients prescribed anti-malarials after introducing RDT guideline at the intervention sites.

There was no significant reduction in prescribing of anti-malarials at any of the control sites during the corresponding periods.

A sub-group analysis (Table 28) shows that the decline in usage of anti-malarials at the intervention sites was more notable among older participants (≥5 years old) than in younger children (<5 years old). The decline among participants who were ≥5 years old versus those who were <5 years old was 5.5 times higher in Tororo (high transmission), 1.5 times higher in Jinja (medium transmission) and 1.2 times higher in Mubende (low transmission).

Table 28: Change in proportion of patients prescribed antimalarials at intervention HCs, by age group

		< 5 years			< 5 years		
District	**Setting**	N*	%	(95% CI)	N*	%	(95% CI)
Tororo	High	1804	-7.7	(-7.73, -7.67)	1469	-42.1	(-42.13, -42.07)
Jinja	Medium	605	-36.7	(-36.79, -36.61)	2015	-55.8	(-55.83, -55.77)
Mubende	Low	343	-52.0	(-52.06, -51.94)	1201	-73.3	(-73.31, -73.29)

*N: total number of febrile participants analysed at each site, before and after intervention. Summary statistics showing number analysed and number prescribed antimalarials before and after intervention is shown in Appendix 12

4.3.5 Change in prescribing of antibiotics after intervention

Table 29 below shows the change in the proportion of fever patients prescribed antibiotics at intervention sites after introducing RDT-based guidelines. Corresponding figures are presented for relevant control sites. Data showing the number prescribed antibiotics before and after introducing RDT-based guidelines are found in Appendix 13

Table 29: Change in proportion of patients prescribed antibiotics after intervention

		Intervention HCs			Control HCs		
District	**Setting**	N*	%	(95% CI)	N*	%	(95% CI)
Tororo	High	3273	-5.0	(-8.7, -1.3)	3355	-0.4	(-3.8, 3.0)
Jinja	Medium	2081	-2.7	(-7.0, 1.5)	2308	-4.7	(-8.8, -0.6)
Mubende	Low	1544	-4.6	(-9.7, 0.5)	1719	-0.6	(-5.2, 4.0)

*N: total number of febrile participants analysed at each site, before and after intervention. Summary statistics showing number analysed and number prescribed antibiotics before and after intervention is shown in Appendix 13

In general, use of the RDT-guideline did not lead to large declines in the amount of antibiotics prescribed at any of the intervention sites (reduced by 5% or less). There was a marginal but significant decline in the amount of antibiotics prescribed at the intervention HC in high transmission area. No statistically significant declines were detected at the intervention sites in medium and low transmission areas. At the control HCS, there was a marginal but significant decrease in prescribing of antibiotics in Jinja (medium transmission area). No significant changes were detected in high and low transmission areas. In general, there was no systematic pattern in the effects of the intervention on prescribing of antibiotics

4.3.6 Change in cost after introducing RDT-based guidelines

Table 30 shows the change in the cost of medical consumables per case of fever at intervention sites, after introducing RDT-based guidelines. It also shows changes in cost at corresponding intervention HCs. The costs consist of the values of biomedical consumables used in diagnosing and treating a case of fever. The absolute values from which these differences were computed are shown in Appendix 14.

Table 30: Change in average healthcare cost after intervention (2008 US$)

		Intervention HCs		Control HCs	
District	Transmission setting	US$ (2008)	(95% CI)	US$ (2008)	(95% CI)
Tororo	High	0.02	(-0.21, 0.25)	-0.02	(-0.21, 0.25)
Jinja	Medium	-0.33	(-0.54, -0.12)	-0.22	(-0.40, -0.06)
Mubende	Low	-0.50	(-0.69, -0.31)	-0.16	(-0.39, 0.07)

After introducing RDT-supported guideline, no significant change in healthcare cost was detected at the intervention HC in high transmission setting. In medium and low transmission settings, the cost of medical consumables per case of fever fell significantly, by US$ 0.33 (17.7%) and US$ 0.50 (24.5%) respectively. Therefore, the lower the malaria transmission intensity in an area, the larger was the amount of cost-saving which was accrued by the use of RDT-based guidelines.

No significant change in cost was detected at the control HCs in high and low transmission settings. However, a significant decline in cost was detected at the control HC in medium transmission zone, although the amount of change at the corresponding intervention HC was 1.5 times higher.

A sub-group analysis by age (Table 31) shows that among children (<5 years), use of the RDT guideline was associated with an increase in cost at all the intervention HCs—although the increase was not statistically significant in medium transmission setting.

Table 31: Change in healthcare costs at intervention sites, by age groups

		< 5 years		5 + years	
District	Transmission setting	US$ (2008)	(95% CI)	US$ (2008)	(95% CI)
Tororo	High	0.61	(0.47, 0.75)	-0.23	(-0.56, 0.10)
Jinja	Medium	0.47	(-0.09, 0.85)	-0.47	(-0.69, -0.25)
Mubende	Low	0.42	(0.13, 0.71)	-0.65	(-0.85, -0.45)

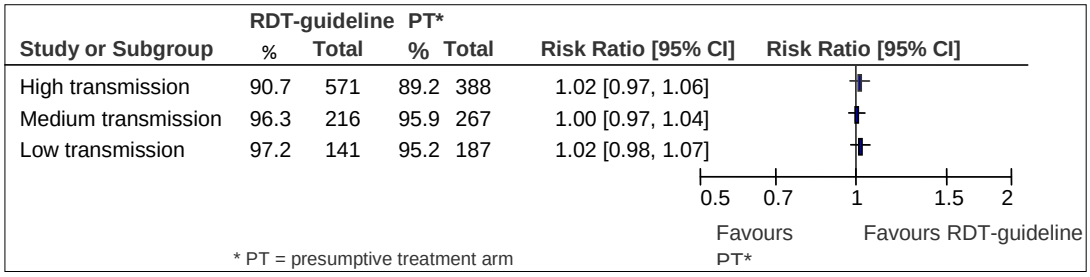
On the other hand, the intervention was associated with significant cost savings among older children and adults (5+ years) at the intervention sites in low and medium transmission settings. The cost of diagnosing and treating an older case of fever (5+ years old) fell by US$ 0.47 (20.7%) in medium transmission setting and by US$ 0.65 (27.1%) in low transmission setting. No significant decline in cost was detected among the older participants in high transmission setting. (Appendix 15 shows the average cost at the intervention HCs before and after introducing RDTs)

4.3.7 Difference in clinical status between comparison groups

During the follow-up survey, a sub-sample (25%) of all fever patients were randomly selected from each intervention and control centre. They were followed-up and their clinical status evaluated at day 5 after treatment initiation. Because assessment of clinical status was not carried out at baseline, comparison of clinical status is made between the pair of HCs within each transmission setting using data from follow-up survey only (data set (d)).

Clinical assessments were based on self-reports, by the patients or their caretakers. Figure 25 below shows the percent of patients from each arm reporting improvement in symptoms after 5 days of treatment. Appendix 16 presents the various perceived clinical states at day 5 by age groups.

Figure 25: Patients reporting improvement in symptoms at day 5 of treatment



A high proportion of fever patients from both arms reported improvement in their clinical status at day 5 of treatment at each of the study sites (89.2% to 97.2%). At all of the transmission settings, no significant difference was detected between the 2 study arms in the percent of fever patients reporting improvement in their symptoms.

Therefore, neither RDT-supported guideline nor presumptive treatment was significantly superior to the other in terms of the effect of their use on clinical outcomes.

4.3.8 Missed malaria cases

Table 32 shows the number of true malaria cases (with positive slides) that were not prescribed any anti-malarial treatment at intervention sites after introducing RDT guidelines. It also shows the false negative error rates due to RDT at corresponding sites.

Table 32: Slide positive cases not prescribed anti-malarials at intervention HCs

Variable	Site	Setting*	N**	%	(95% CI)
Patients with positive BS not prescribed anti-malarials	Tororo	High	1189	1.9	(1.1, 2.7)
	Jinja	Medium	358	5.9	(3.5, 8.3)
	Mubende	Low	174	17.2	(11.6, 22.8)

** malaria transmission intensity in the region*

***N: total number of febrile patients with positive reference slides*

The proportion of slide positive cases that were not prescribed antimalarials in Tororo (high transmission) and Jinja (medium transmission) was 2% and 4.5% respectively. In Mubende (low transmission) more than 17% of slide positive cases were not prescribed anti-malarial treatment. Most (>70%) of the malarial cases that missed anti-malarial treatment were older children and adults (>5 years old) (not shown in Table 21).

Table 32 also shows that the percent of missed malaria cases corresponds with the RDT false negative error rates at the corresponding sites, namely 2%, 4.5%and 17.2% respectively (see Table 26). This implies that the observed under-treatment of malaria at the intervention HCs can be attributed almost exclusively to the inaccuracy of the RDT, rather than to clinician non-compliance.

4.3.9 Slide-negative fevers prescribed anti-malarials

Table 33 shows the number of slide negative cases that were prescribed anti-malarials at intervention sites after introducing RDT guidelines. In addition, it displays the false positive (α) error rates due to RDT at corresponding sites.

Table 33: Slide negative cases prescribed anti-malarials at intervention HCs

Variable	Site	Setting*	N	%	(95% CI)
Patients with negative BS and prescribed anti-malarials	Tororo	High	1032	43.1	(40.1, 46.1)
	Jinja	Medium	622	15.0	(12.2, 17.8)
	Mubende	Low	419	11.0	(8.0, 14.0)

** malaria transmission intensity in the region*

***N: total number of febrile patients with negative reference slides*

The results show that, during RDT era, the proportion of cases prescribed anti-malarials at high, medium and low transmission sites was 43.1%, 15.0% and 10.5% respectively. These numbers correspond with the false positive error rates due to RDT at the respective intervention sites, which are 43.9%, 15.2% and 10.5% respectively (see Table 26). Therefore, use of the guideline did not eliminate over-use of antimalarials totally. This over-use of antimalarials can be attributed almost entirely to the diagnostic inaccuracy of RDT rather than to clinician non-compliance.

4.3.10 Summary of main findings

The analysis shows a high level of clinician adherence to RDT-supported treatment. At least 97% of all fever patients were tested for malaria with RDT across all intervention sites. In addition, nearly all (>98%) patients testing positive on RDT were prescribed anti-malarials. On the other hand, practically none of those testing negative were prescribed any anti-malarials.

Prior to introducing RDT-based guidelines, the majority (>90%) of fever patients were prescribed anti-malarials. After introducing RDT-supported guidelines, anti-malarial prescribing declined by 22%, 49% and 60% at the intervention sites in settings with high, medium and low malaria transmission respectively. Relative to younger patients (<5 years), anti-malarial prescribing declined in older patients (≥5 years) at the corresponding sites by a factor of 5.5, 1.5 and 1.2 respectively.

Despite the near-perfect compliance with RDT-based guideline, their use did not completely eliminate over-use of anti-malarials at intervention HCs. The percent of slide-negative patients prescribed anti-malarials in the era of RDT was 43.1%, 15% and 11% in high, medium and low transmission setting, corresponding with the RDT α-error rates at the respective sites.

RDT-supported treatment resulted in some malarial (slide positive) fever patients not receiving anti-malarials at the intervention sites. The proportion of malarial fever patients not given anti-malarials at the intervention sites at high, medium and low transmission sites was 1.9%, 5.9% and 17.2% respectively. These figures closely correspond with the RDT false-negative error rates at the respective sites, namely 2.0%, 4.5% and 17.2% respectively.

Use of the RDT-based guideline did not have any clear-cut effect on antibiotics usage.

The analysis did not demonstrate convincingly whether or not the use of RDT was superior to presumptive treatment in terms of patients' clinical outcomes. Reported improvements in symptoms were similar (>90%) in both arms, but the findings were not statistically significant.

Prior to the intervention, the cost of treating a case of fever presumptively ranged from US$1.45 (high transmission) to US$ 2.0 (low transmission), with no imbalance between

any pair of comparison HCs. After the intervention, no significant change in cost was detected at the intervention HC in high transmission setting. At the HCs in medium and low transmission settings, the post-intervention cost of diagnosis and treatment fell by US$ 0.33 (17.7%) and US$ 0.50 (24.5%) respectively. In both medium and low transmission settings, cost-savings were accrued exclusively in older children and adults: US$0.47 (20.7%) and US$0.65 (27.1%) respectively. Use of the guidelines in younger children resulted in additional costs to the healthcare system at all the intervention sites.

4.4 Discussion

This chapter uses a dataset from a pragmatic trial from Uganda to answer the question whether an RDT-based guideline can be sufficiently implemented in routine practice and whether its application in routine practice can lead to better prescribing, better clinical outcomes and significant cost-saving. Effects are compared relative to a treatment policy for fever based on clinical judgement.

4.4.1 Quality of implementation of guideline

a) Clinician adherence

The results of this analysis show that, even in routine practice settings, HWs can request RDTs and respond to the results as per the new guidelines They suggest that HWs may have high confidence in RDT-based guidelines in routine practice. These findings contrast with the results observed in 3 of the 4 trials synthesized in chapter 3 in which as many as 40% to 80% of patients with negative RDT results were prescribed antimalarials [34-36]. They also differ from evidence from observational studies carried out in routine clinical settings, which show that large numbers of fever patients are either treated presumptively or are prescribed antimalarials despite having a negative RDT results—despite exposure to training, and the presence of RDTs and job aids [32, 33]. However, these findings are consistent with those from Zambia by Yeboah and colleagues where HWs used RDTs and prescribed antimalarials according to the results most of the time. A more recent non-randomised study from Uganda—carried out in the same epidemiological settings as this trial—detected high utilisation of RDTs. However, an average of 30% of fever patients with negative RDTs were given antimalarials despite regular support supervision [11], which also found high RDT utilisation by health workers. Therefore, the adherence levels shown in this analysis appear to be uncommon.

It is unclear why HWs showed high level of adherence to RDT results in this trial. Literature suggests that factors such as being a paraprofessional staff [90, 112, 146], adequate in-service training [90, 108, 117, 146] and support supervision [93, 112, 117, 146] are associated with high level of adherence to guidelines. The cadres of staff involved in this trial appear to be similar to those in the trials analysed in chapter 3 and in several other observational studies. In this trial, there was no formal support supervision offered to HWs once data collection has commenced. Therefore, in this

trial, high utilisation of RDTs and adherence to results may be attributed to adequate training and the continued informal interactions with researchers. However, due to data limitation, we do not know how proficient HWs were in implementation other components of the guideline, notably clinical assessment, the multiple steps in carrying out an RDT test, and patient counselling. These limitations are discussed further in section 4.4.7 (a).

b) Accuracy of RDT

In addition, the analysis confirms that use of HRP-2 RDTs is characterised by varying levels of false positivity, which is considerably high in high transmission settings. In this analysis, over-use of antimalarials remained substantially high in high transmission settings. This indicates that, despite adequate clinician compliance with the guideline, use of RDT-based guidelines does not totally eliminate over-use of anti-malarials—especially in high transmission settings. This undermines a key objective of RDT-based policies in high malaria transmission areas: which is to reduce over-use of antimalarials and to save cost [14, 32-35, 53].

The evidence from this analysis and other studies [135, 161] suggest that use of RDTs to guide treatment of fever may be associated with a risk of malarial fevers not being prescribed antimalarials, due to the lower sensitivity of RDT relative to presumptive treatment of fever. In this analysis, this risk was highest in a low transmission setting, where the sensitivity of an HRP-2 type RDT was only 82.8%. The sensitivity of HRP-2 RDTs falls at parasite concentration < 50 parasites/μL of blood, e.g. in non-immune population in low prevalence areas and in children [22] (see literature review, section 2.8.3). Therefore, in non-immune individuals, clinical disease can occur at parasite densities lower than the lower detection limits of the common RDTs (see literature review). In this analysis, it is unclear if the high false negative error rate in low transmission setting is the result of errors committed in carrying out the tests, or if it is due to reduced sensitivity of RDTs at low parasite concentrations.

Two previous evaluations of HRP2-based assays were conducted in 1999 (*Para*Sight™-F test)[162] and 2002 (*Paracheck Pf*® test, Orchid Biomedical Systems, Goa, India)[163] in malaria transmission settings similar to Mubende. In both studies, RDTs were performed by health unit staff and results compared against reference microscopy. In both studies, the RDTs evaluated had sensitivity of at least 97% at parasite density

above 500/μL, and specificity of at least 86 % against. In one of the studies, sensitivity was reportedly lower (figures not provided) at parasite densities below 100/ μL [163]. A more recent assessment of RDT performance in similar settings, carried out during both dry and rainy seasons, detected an average sensitivity of 91% of an HRP2-based assay (*Paracheck Pf*® test, Orchid Biomedical Systems, Goa, India) [164]. Therefore, a sensitivity of 82.8% for an HRP2-based assay (*Paracheck Pf*® test (Orchid Biomedical Systems, Goa, India) appears atypical, and might signal a lack of proficiency in the use of RDTs by health unit staff.

An objective of introducing RDT-based policies instead of presumptive treatment is to target antimalarials to slide positive fever cases [19, 138]. These results suggest that RDT guided treatment of fever may not result in better targeting of antimalarials in slide positive cases, as some of them are missed. Presumptive treatment of fever is more likely to cover all or most malarial fevers with antimalarials than RDT-guided treatment.

4.4.2 Effect on prescribing of anti-malarials

The analysis also suggests that, despite the high false positive error rate associated with HRP-2 type of RDTs, use of the guideline in routine clinical settings can considerably reduce anti-malarials prescribing at all transmission settings. However, use of the guideline is more effective in reducing usage of antimalarials in older participants (5+ years) than in children (<5 years), and in lower transmission settings than in high transmission settings.

In chapter 3, it was noted that the impact of RDT-based guidelines on prescribing of antimalarials is dependent on the degree of HW compliance with RDT results. For example, usage of antimalarials declined by 77% point (RR = 0.23; 95% CI: 0.14, 0.38) in a trial where HWs prescribed antimalarials to RDT-negative patients 0.6% of the time [37]. On the other hand there was no difference in prescribing of antimalarials between RDT and presumptive treatment arms in trials where HWs prescribed antimalarials to RDT-negative patients more than 80% of the time (RR = 1.02, 0.95, 1.09) [34]. Trials in which HWs showed partial compliance detected marginal declines in prescribing of antimalarials [35, 36]. Evidence from this chapter shows the effect of RDT-based policies in a context of near perfect compliance. Therefore, in the pragmatic trial, variability in prescribing across sites is explained largely by the local malaria prevalence. This clearly illustrates that local malaria epidemiology is an important determinant of the

magnitude of effect of RDT-based policy on prescribing of antimalarials [14, 21, 22, 139, 144]. Implementation of RDT-based policies may need to be targeted according to pattern of malaria prevalence in a country, rather than adopting a one-size-fits-all policy for all regions within the country. The trials synthesized in chapter 3 did not disaggregate results by malaria prevalence—which undermines our ability to compare the results of this analysis with those in chapter 3.

4.4.3 Effects on prescribing of antibiotics

RDT-based policies are intended to aid the identification of parasite-negative individuals in whom alternative diagnoses can be sought [19, 70-72]; and to aid better targeting of antibiotics relative to clinical diagnosis [72, 165]. Owing to the lack of capacity to differentiate between the different causes of non-malarial fevers at lower level health facilities, use of RDT-based guidelines could lead to increased and unnecessary use of antibiotics [138, 155].

The results of this analysis suggest that use of RDT-based guidelines may not have significant effect on antibiotics usage, especially in low and medium transmission settings. A small (4.7%) but statistically significant reduction in prescribing of antibiotics was also detected at the control HC in Jinja (medium) transmission. The reason for this is unclear; perhaps it was due to shortage of antibiotics. In the absence of definitive diagnoses of the causes of non-malarial fevers, this analysis cannot judge if prescribing of antibiotics for RDT-negative fevers, or if withholding them from RDT-positive cases constituted rational prescribing [138].

In chapter 3, it was noted that one of the trials detected no significant difference in prescribing of antibiotics between the two comparison arms (RR = 0.97, 95% CI: 0.90 to 1.05)[34]; while one trial detected a significant (19%) increase in prescribing of antibiotics in the intervention arm (RR = 1.19, 95% CI: 1.06 to 1.34), especially in older age group (>5 years) [36]. However, a cross-over trial from Zanzibar (Tanzania) found that use of RDT-supported guidelines was associated with a 40% increase in antibiotics prescribing [121]. A more recent observational study in Tanzania also found that, relative to presumptive treatment of fever patients, use of RDT-based guidelines was associated with a 44% increase in antibiotics usage [156].

Therefore, the evidence from this analysis and previous studies indicate that use of RDT-based guidelines in routine clinical settings may not result in a significant change in antibiotics prescribing. In any case, it may lead to an increase in prescribing of antibiotics.

4.4.4 Clinical outcomes

One of the drawbacks of RDT-based policies relative to clinical diagnosis is the increased risk of mortality and morbidity from missed malaria cases, especially in children [26]. This analysis detected no statistically significant difference in clinical outcomes among fever patients treated according to RDT guidelines versus those treated clinically. This suggests that the low sensitivity (17.2% false positivity error rate) of RDTs at the low transmission site did not influence morbidity significantly.

We noted in chapter 3 that the two trials which evaluated effect of the intervention on clinical outcomes also detected no significant difference in clinical outcomes between the comparison arms (RDT vs. clinical diagnosis) [34, 37]. However, the cross-over trial from Zanzibar (Tanzania) detected a statistically significant reduction (from about 5% to 2.5%) in the risk of persistent symptoms after introducing RDT guidelines (OR = 0.5, 95% CI: 0.3 to 0.9, p = 0.005)[121].

In general, the evidence from this analysis and those in chapter 3 suggests that neither an RDT-supported policy nor a presumptive treatment of fever is significantly superior to the other in terms of their effects on clinical outcomes. However, it is possible that the analyses in chapters 3 and 4 failed to detect a difference between the comparison interventions because of design and measurement limitations.

We know from literature that measures of clinical outcomes are usually not sensitive enough for demonstrating differences in quality of care, especially subjective ones such as those evaluated in this thesis (chapters 3 and 4) [85]. The intervention is applied at the level of HWs, yet clinical outcomes are assessed at the level of the patient. Clinical outcomes may be influenced by several factors other than the intervention—such as effectiveness of treatment, patient compliance with treatment, and demographics [85]. As such, clear differences in the quality of care as shown by process measures may not translate into differences in clinical outcomes [85]. Therefore, a lack of improvement in clinical outcomes (in both chapters 3 and 4) may indicate that the indicators of clinical

outcomes evaluated in chapters 3 and 4 are not valid for demonstrating differences in the effects of the interventions; and may not imply that there was lack of improvement in the overall quality of care [85].

Further, different components of a complex intervention (such as the guideline under evaluation) may be linked to different outcome indicators [74]. For example, the value of antimalarials saved per patient is directly linked to the extent to which HWs use RDTs and prescribe antimalarials relative to RDT results. On the other hand, improvement in clinical outcomes is expected mainly from improved management of non-malarials fevers [19, 70-72]—as long as HWs are proficient in using RDTs. Management of RDT negative patients at low referral facilities requires good clinical acumen which a dispensary or HC staff might lack [138]; and adequate clinical assessment [138, 155, 157, 166]—which we have not been able to evaluate in both chapters 3 and 4 because of data limitation. It is possible that there was no difference between the comparison interventions in the quality of care offered to parasite negative patients.

4.4.5 Treatment costs

This analysis suggests that use of RDT-supported policy instead of presumptive treatment of fever can reduce healthcare cost in normal clinical practice by 25% and by 18% in low and medium malaria transmission settings respectively. The potential of the policy to save cost in high transmission settings is questionable. More notably, significant savings are accrued exclusively in older patients (5+ years). In children (< 5 years old) use of the policy to guide treatment of fever results in added cost to the healthcare system in all transmission settings, despite adequate clinician adherence to guidelines.

A significant decline in cost worth US$ 0.22 (95% CI: US$-0.40 to US$-0.06) was detected at the control HC in medium transmission setting (the cost-saving at the corresponding intervention site was US$ 0.33). This may be a signal that declines in cost (or resource use in general) may be a result of shortages of supplies or inputs in the health system—a factor which could not be investigated by this analysis.

In this particular analysis which is based on HRP-2 type of RDT, the potentials for cost-saving at the high transmission site could have been partly undermined by the low specificity of RDTs in this setting [138].

The analysis in this chapter shows the amount of cost that can be saved in different malaria transmission settings and in different age groups when HWs comply with an RDT-based guideline adequately. These findings are consistent with those from the economic models summarised in the literature review, which suggest that use of RDTs to guide fever management can reduce healthcare cost by 21% to 25% in low to moderate malaria prevalence areas if clinicians adhere adequately with RDT-based guidelines [12, 13, 33, 72, 140, 144]. The findings of this analysis contrasts with those from another pragamatic trial from Tanzania. Whereas use of RDTs was associated with a 21% reduction in healthcare cost in moderate malaria transmission areas, its use was associated with a 41% increase in cost in high transmission areas [33]. The reason for this was that HWs did not follow the guideline in low transmission areas [33].

Therefore, when applied in a context of adequate HW adherence to guidelines, RDT-supported policy for malaria can save money in routine practice in low and medium transmission, especially if their use is limited to older children and adults. In high transmission settings, the effect of the policy on cost-saving is questionable, even among older patients—even if HWs adhere adequately with the guideline.

It can be noted that in Jinja (medium transmission), the proportion of participants with fever was consistently higher in the control HC than in the intervention HC, despite the fact that the pair of HCs were similar in staffing, participants' age and (lack of) capacity for microscopy. Because analyses of resource use were based on before-and-after comparisons, the imbalance between comparison HCs may not have significantly influenced the results at the respective HCs. However, the imbalance may have implications for budget impact analysis since we do not know which of the pair represents the typical HC in the region in terms of epidemiology of fever.

4.4.6 Usefulness of the analysis

The findings of this analysis can inform decisions as to whether it is worthwhile, from an economic perspective, to scale-up the use of RDT-supported guideline to eligible HCs in all transmission settings, or whether to limit its application to low and medium

transmission settings only; and whether to use it in all age-groups or target older children and adults (5+ years) only [26, 70, 138].

These results suggest that it makes economic sense to restrict the use of RDT-based policy to older patients (5+ years) and to low and medium transmission zones only. Its use in children adds cost to the healthcare system, while the potential for cost saving in high transmission areas is equivocal. However, varying diagnostic strategies for malaria by transmission settings can pose practical challenges in implementation [26, 70]. Therefore, from a practical standpoint, the policy could be implemented across the whole country as long as its use is limited to 5+ years old. Cost-effectiveness analyses in which longer-term and broader outcome measures such as the DALY are used, suggest that use of RDT-based guideline can be cost-effective (95% certainty) at malaria prevalence of up to 62% [72, 140, 144].

Thus, from practical and economic standpoints, it seems sensible to scale-up the use of RDT policy to HCs without capacity for microscopy in all malaria transmission settings in Uganda, as long as its use is limited to older patients (5+ years).

This analysis offers a comprehensive evidence base on the merits of RDT-based policy (accuracy of RDT, clinician adherence, prescribing practices, outcomes and cost implications) based on patient-level data. Therefore, it offers a relatively more accurate and more complete basis for decision making in Uganda, than the existing body of evidence reviewed in this thesis. In addition, the results from this analysis provides more accurate data which can be used in modelling cost-effectiveness of RDT-based policy in Uganda, and similar contexts.

4.4.7 Limitations of the study

a) Assessing the quality of implementation

The RDT guideline for Uganda describes 5 major components of the healthcare process that needs improvement: (a) assessment, which includes medical history and clinical examination; (b) use of RDT; (c) patient classification (diagnosis) (d) treatment and (e) counselling [157]. In this chapter, analysis of HW adherence is based on the use of RDTs and prescribing of anti-malarials relative to RDT positive and RDT negative results only. Use of these 3 indicators alone in the assessment of adherence to the guideline fails to fully describe the quality of the diagnostic processes in the care of patients with fever, and the degree of adherence to the guideline [109, 113, 149]. In particular, the quality of clinical assessment is crucial in the differential diagnosis of the causes of RDT-negative fevers, and in targeting antibiotics. Therefore, a comprehensive and simultaneous assessment of all components of the guideline is necessary in order to identify all the elements which need improvement. This limitation is addressed by the assessment described in Chapter 5.

b) Range of inputs included in cost analysis

In this analysis, I included the cost of biomedical consumables only. Other programmatic costs (e.g. staff time, the cost of training and support supervision) were excluded because they were not contained in the database, and because the key informants were not available to provide the necessary information. The cost implications of an intervention depend on the types and range of inputs included in the analysis [72, 140]. If these programmatic costs were also included the analysis, the results could have shown a smaller amount of cost savings attributable to RDT-supported treatment, even in low and moderate transmission areas; or the analysis could have shown added cost to the healthcare system instead [72, 140]. On the other hand, the analysis only captures short-term outcomes, which fails to reflect all the possible consequences of treatment. When health outcomes are measured in terms of longer-term and broader measures such as the DALY, use of RDT-based guideline is likely to be cost-effective (95% certainty) at malaria prevalence of up to 62% [72, 140, 144]. Therefore, this analysis does not capture the full economic worth of RDT-based policy. Nevertheless, the evidence base provided by this analysis is sufficient to answer the policy-relevant questions that this analysis set out to answer. The evidence from this

analysis can feed into future cost-effectiveness analysis of RDT-based policy in Uganda or similar countries.

c) Skewed cost data

The analysis uses cost data that is skewed to the right. As such the mean cost values shown in this analysis are slightly higher than the corresponding median values. This is typical of trial-based cost data [143]. The standard approach would have been to calculate costs in terms of medians [143]. In terms of costs, this is inappropriate because the decision maker needs to be able to link the summary measure of cost per patient with the overall budget impact [143]. This can only be achieved with the mean. Methods of dealing with skewed cost data is still an area for research [143]. Those that have been suggested—such as data transformations and non-parametric methods—have limitations and are contested [143, 167, 168]. Because of these methodological uncertainties, and because of the need to use summary measures which can be used for budget impact analysis, costs were presented in form of the mean.

d) Lack of data on malaria status at baseline and at control HCs

The dataset does not include data on malaria status of participants (based on blood slides) at baseline, and at control HCs. Therefore, this analysis could not estimate the degree to which RDT-supported treatment guideline might have reduced overuse of anti-malarials.

4.5 Conclusion

This thesis has demonstrated that RDT-based guidelines can be implemented sufficiently in routine practice. It has also shown that when applied in a context of adequate HW adherence, use of RDT-supported guidelines for fever can reduce prescribing of antimalarials and save money in primary care settings in low and medium transmission settings in Uganda, especially if their use is limited to older children and adults. In high transmission settings, the effect of the policy on cost-saving is questionable, even among older patients. This thesis indicates that use of RDT-based guideline may not significantly change prescribing of antibiotics in routine practice, although evidence from other pragmatic trials suggest that its use could lead to a significant increase in prescribing of antibiotics in routine practice. Although its use is associated with increased risk of missed malaria cases, the effect of the policy on clinical outcomes is uncertain. More evidence from randomised trials is required to further examine effects of RDT-based policies on clinical outcomes relative to presumptive treatment in routine clinical settings.

Therefore, the policy for treating fever based on RDTs instead of clinical diagnosis may significantly avert irrational usage of antimalarials and save healthcare cost in routine practice. However, its effect on clinical outcomes in routine practice is uncertain. If the decision to scale-up implementation of RDT-based guidelines is based purely on economic considerations, then the analysis in this chapter and previous models suggest that it would be sensible to restrict implementation of the policy to low and medium transmission settings, and to use it in older patients (5+ years old) only. Considering the practical challenges associated with implementing a diverse policy for a single disease in a country, it seems more prudent to scale-up the use of RDT policy to eligible HCs in all malaria transmission settings in Uganda, as long as its use is limited to older patients (5+ years).

Chapter 5

Survey of current practice

Chapter 5 Implementation of RDT-based guideline in current practice

5.1 Introduction

5.1.1 Overview

This chapter consists of a survey of current practice to evaluate the quality of actual practice and adequacy of support services in the districts where RDT-based guidelines have been rolled out for small-scale implementation in Uganda. The survey uses a tool which consists of a checklist for assessing the quality of the following components of the guideline: (a) clinical assessment, (b) use of RDTs, (c) patient classification, (d) treatment prescribed, and (e) advice on medications. Additionally, the tool includes checklists for assessing adequacy of essential supplies, training and supervision. The survey assesses whether HWs can sufficiently implement all components of an RDT-based guideline in routine clinical settings in Uganda. Assessment is carried out in children (< 5 years old), and in settings with high and low malaria transmission respectively. In addition, it assesses the adequacy of inputs in clinical practice and of support activities in the selected districts. The survey focused on children (<5 years) mainly because the tool used in this study was adapted from an HFA tool designed for assessing implementation of Integrated Management of Childhood Illness (IMCI) guidelines [169, 170]. The generic HFA tool has been field-tested before [170] and used in several previous surveys [92, 171, 172]. Evidence on the relationship between age of study participants and HW adherence to guidelines (prescribing) is generally equivocal [90, 91, 93, 108, 112, 115, 116].

5.1.2 The health system in Uganda

a) Service delivery

Health service delivery system in Uganda is structured into 7 referral levels, namely HCI, HCII, HCIII, HCIV, general hospital, regional referral hospitals and national referral hospitals. Table 34 below shows the designated geographic coverage, the target population and the services provided at each level of facility. Appendix 2 summarises the recommended staffing for each level of care.

Table 34: Levels of health service delivery in Uganda (2011)

Responsibility	Referral level	Administrative level	Target population	Services provided
District local government (district health system)	HCI	Village	1,000	Community-based preventive and promotive health services
	HCII	Parish	5,000	Preventive, promotive, outpatient curative services, outreach care
	HCIII	Sub-county	20,000 – 25,000	All services at HCII, maternity, inpatient services and laboratory services
	HCIV	County	100,000	All services at HCIII, emergency surgery and blood transfusion
	General hospital	District	300,000 – 500,000	All services at HCIV, radiology, in-service training, consultation & research to community-based programmes
Ministry of health	Regional referral hospital	Regional	2,000,000	All services at general hospitals, specialist services: pathology, psychiatry, ENT, ophthalmology, dentistry, intensive care, consultant surgical and medical services
	National referral hospital	National	Country wide	Specialist services, teaching and research

Source: Health Sector Strategic Plans (HSSP) II and III (Uganda) [173, 174]

Basic health services are provided through HCIs to general hospitals. The responsibility for the delivery of basic services is devolved to the district local government. Specialist services are offered at regional and national referral hospitals which are autonomous. The Ministry of Health (MoH) is responsible for the formulation and dissemination of policies, strategic plans, standards, guidelines; and for monitoring health sector performance. In addition, the MoH is responsible for nationally coordinated programmes such as emergency preparedness and epidemics control [173, 174].

An HCI consists of a team of community volunteers (the Village Health Team (VHT)) that works as a link between health facilities and the community. It has no physical structure. HCIIs are the first level of contact between the formal health sector and the communities. They serve the majority of the population. HCIIs provide out patient care and community outreach services only. They have no capacity for microscopy. An HCII is manned by 2 or 3 staff consisting of an enrolled nurse, enrolled midwife and/or a nursing assistant [173, 174].

The new RDT-based guideline for fever is intended to be used at HCIIs, which lack the capacity for microscopy, and where the most qualified staff is either an enrolled nurse or enrolled midwife.

b) RDT-based guideline for Uganda

In Uganda the RDT-supported policy for treating fever was adopted in 2009. The Uganda Ministry of Health has accordingly revised its guidelines and training manual in line with WHO recommendations [3, 157]. RDT-based guidelines have since been introduced in government-owned Health Centres (HCs) without the capacity for quality microscopy. So far only 6 districts are covered, out of a total of 72:—4 from a low malaria transmission zone and 2 from a high transmission zone.

The new guideline has 7 main components [157]:

1) Assessing patients with fever and selecting patients for RDT testing
2) Performing and reading an RDT
3) Managing a patient with fever and a positive RDT
4) Managing a patient with fever and a negative RDT
5) Recognition and referral of patients with severe illness
6) Patient education
7) RDT storage and monitoring

Although the new RDT-based guideline came to effect only recently, ACT has been in use as the first line treatment for malaria in Uganda since 2002. The new RDT-based guideline incorporates the previous guideline on the use of ACTs [175]. In addition to specifying when to prescribe ACT and other antimalarials, the new guideline also describes the dosages of these drugs by age-group and weight[8] [157].

c) Supply chain in the public sector

Health facilities in the public sector receives supplies through two distribution mechanism: pull and push systems

Pull system

[8] The dosing of the recommended ACTs is more complex than that of the anti-malaria drugs previously used (CQ-SP combination) 157. Uganda Ministry of Health Malaria Control Programme, *User's Manual: Use of rapid diagnostic tests (RDTs) for malaria in fever case management in Uganda; Near-final draft: 31 January 2009*, 2009.

In the main, drugs and medical supplies are procured and distributed to health facilities through a pull (demand-based) system. In this system, individual health units determine the types and quantities of drugs and supplies required based on local demand patterns, distribution frequencies, costs and inventory level [176-178]. Typically, individual health unit orders are consolidated into one order at Health Sub-District (HSD) level. Each HSD submits its consolidated order to the district, which submits them separately to National Medical Stores (NMS)—a government parastatal responsible for procuring and supplying medical goods within the public system. NMS packs the order for each HSD, which is then invoiced on behalf of the health units. The HSD then unpacks the consolidated NMS consignment and re-packs and distributes them according to original health unit orders. Occasionally, health unit orders are individually packed and invoiced by NMS, but distributed in one consignment to the district. They are then distributed to individual health units as described above.

Push system

Occasionally selected items are supplied through the push (kit) system. A central authority (e.g. NMS, an NGO) determines the types and quantities of items required by health units. Health units of the same referral level receive a standard kit (same items, in the quantities) of the selected items [176-178]. Health unit orders are individually packed at NMS but are distributed in one consignment to the district headquarters. Supplies are unpacked, re-packed and then distributed to individual health units by the district supplies officers or by HSD heads. Cost of procurement and distribution is borne by the central authority. The push system is normally used in emergency relief efforts where demand exceeds supply, e.g. insecurity, outbreaks, natural disaster [176]. The push system also applies where the health system cannot satisfy normal demand patterns for reasons of cost and availability, e.g. when an expensive product has just been introduced into the system and/or is not readily available in the country [176].

Supply of ACTs and RDTs

At the time of the survey, ACTS were provided to health facilities mainly through the pull system while RDT kits were provided through the push system. Since the official introduction of the RDT-based guideline, supplies of both ACTs and RDTs have been erratic and unreliable. At the time of this survey , all eligible HCs had run out of RDT kits for a period of 8 to 9 months, forcing clinicians to revert to presumptive treatment.

As part of this survey, additional stocks of RDT kits and ACTs were supplied to selected HCs through the district supply chain system prior to the survey dates..

5.1.3 Aim

To assess the quality of actual practice when an RDT-based guideline is rolled out to health services in a district as a whole

5.1.4 Research questions

Is the RDT-based policy being implemented sufficiently in the districts where it has been rolled-out in Uganda?

5.1.5 Specific Objectives

1. To establish if all components of an RDT-based guideline for fever are implemented to acceptable standards in Uganda

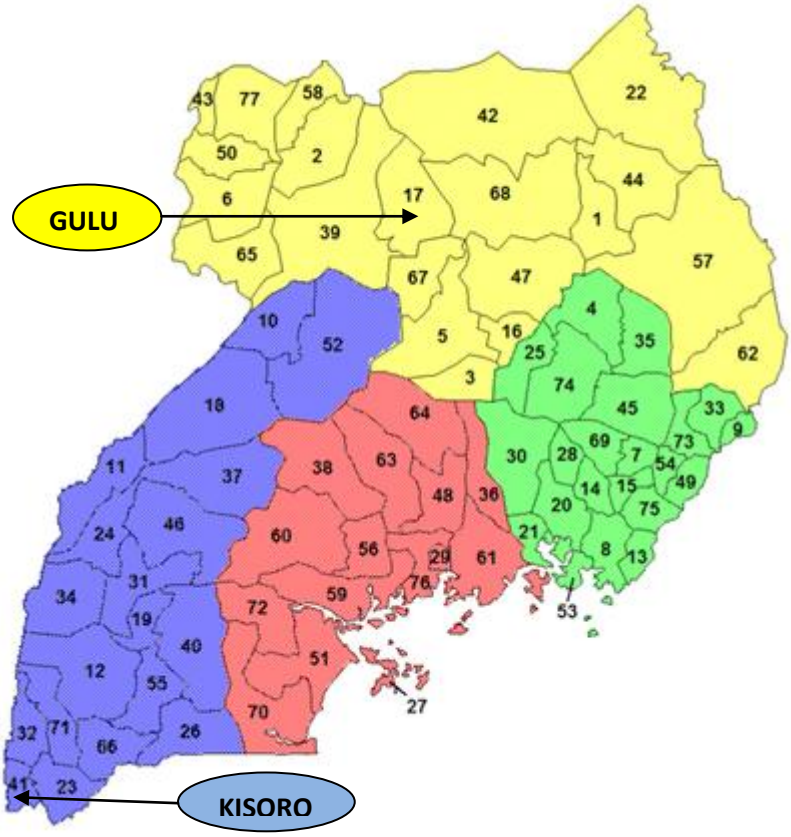
2. To determine if essential inputs and support services required for effective implementation of RDT-guidelines in febrile children are adequate in Uganda

5.2 Methods

5.2.1 Study location

The study takes place in two districts where RDT guideline has been introduced by the Ugandan Ministry of Health (MOH) for small-scale implementation: (a) Kisoro in South West Uganda—a setting with low malaria transmission; and (b) Gulu in Northern Uganda—a setting with high malaria transmission (See Figure 26).

Figure 26: Map of Uganda showing the two districts where the survey was conducted



The two districts are among the six districts in which the Uganda MOH has introduced the guideline during 2009: 2 districts in high malaria transmission area, and 4 districts in low malaria transmission area. The survey took place at government-owned health centres (HCs) of referral level II (HCIIs), where RDT-based guidelines had been introduced. HCIIs represent the first level of clinical care, a setting without the capacity for microscopy.

5.2.2 Design

A survey was carried out in the 2 districts in November and December of 2011. The two districts represent a practical setting for Uganda, which is a context of chronic shortages of medical supplies. In light of foreseen shortages, the study supplied HRP2-based RDT kits (Clearview® Malaria P.f., British Biocell International Ltd) and ACTs (Artemether/Lumefantrine 20/120 mg tablets). These supplies were made available through the normal supply chain to avoid the risk of preferred response bias[179, 180]. Supplies were delivered to the respective district health offices, from where they were distributed to selected facilities by the district supplies officer at least one day prior to the survey date. Health workers were not aware that the study supplied these materials. Therefore, they did not have prior information about the study's purposes and what behaviour was expected of them. Health workers were asked (by the supplies officers) to use the items as they deemed fit.

The survey consisted of 3 main parts: (a) an assessment of HW performance of a set of clinical tasks which are specified in the RDT-based guideline, (b) an assessment of the adequacy of selected inputs (staffing, drugs and equipment) that are essential in the provision of care to a child with fever, and (c) an assessment of the adequacy of support activities (training and support supervision).

The survey used a tool which comprised of a checklist for assessing the quality of the following components of the guideline: (a) clinical assessment, (b) use of RDTs, (c) patient classification, (d) treatment prescribed, and (e) advice on medications. Additionally, the tool included checklists for assessing adequacy of essential inputs, training and supervision. The survey tool was adapted from a Rapid Health Facility Assessment (R-HFA) tool for assessing implementation of guidelines for Integrated Management of Childhood Illness (IMCI) [169, 170]. The tool was modified by incorporating items for assessing specific RDT-oriented procedures specified in the RDT-guideline for Uganda—specifically, the steps in carrying out an RDT and essential equipment and supplies which are required in order to carry out an RDT effectively [157].

The survey uses the Lot-Quality Assurance Sampling (LQAS) method to judge if the quality of practice, or if the quality of essential supplies and support services in the surveyed district is acceptable, basing on pre-defined performance benchmarks. LQAS

principles were discussed in chapter 2. Their application in this survey is described in the following sections.

5.2.3 Sampling and decision rules

A three-stage sampling design was applied. The 2 districts in which the survey was located were randomly selected: 1 from a high malaria transmission zone, and 1 from a low malaria transmission zone. In each zone, a district was selected by randomly drawing a name from a hat containing names of eligible districts in the transmission zone. Then a 2-stage LQAS-based sampling design was applied:

- to select a sample of Health Centres (HCs) within each district to determine whether an acceptable proportion of the HCs met a set performance target for specific indicators measured in each of the aforementioned modules (5.2.2 above)
- to select a sample of consultations within each HC consisting of children (<5 years) with fever, which were observed in order to judge the adequacy of a single provider's performance.

a) Sampling HCs

Sample size

To classify performance of HCs in a district as either high or low, four parameters were set beforehand, in consultation with the Ministry of Health officials.

i. The desired performance threshold (p_U): At least 80% of HCs in each district was expected to demonstrate adequate performance for each specific indicator included in the assessment.
ii. A lower threshold below which performance was deemed highly unacceptable (p_L) was set at 50%.
iii. The probability of misclassifying a district with high performance as having low performance (α error) was set at $\leq$0.10.
iv. The probability of misclassifying a district with low performance as high (β error) was also set at $\leq$0.10.

These parameters were used to calculate the LQAS sample size of HCs to enrol in the study, to obtain the decision rule, and to obtain the exact values of the associated misclassification errors. Table 35 below shows the number of eligible HCs in each

district. Owing to the small total number of HCIIs in each district, the hypergeometric rather than the binomial formula was used to calculate the sample size and to determine the decision rule for each district [181]. The hypergeometric formula provides a finite population correction when computing the sample size for a population [104]. Basing on the conditions above and the number of HCIIs in each district, the hypergeometric formula yielded the sample size (n), decision rules (d) and classification (α and β) errors shown in Table 35 below.

Table 35: Number of HC II, sample size and decision rules

District	Number of HC IIs (N)	Sample Size (n)	Number selected	Decision rule (d)	α-error	β-error
Kisoro	12	8	8	6	0.078	0.030
Gulu	21	10	10	7	0.049	0.095
TOTAL	**33**	**18**	**18**	**-**	**-**	**-**

Decision rule for district performance

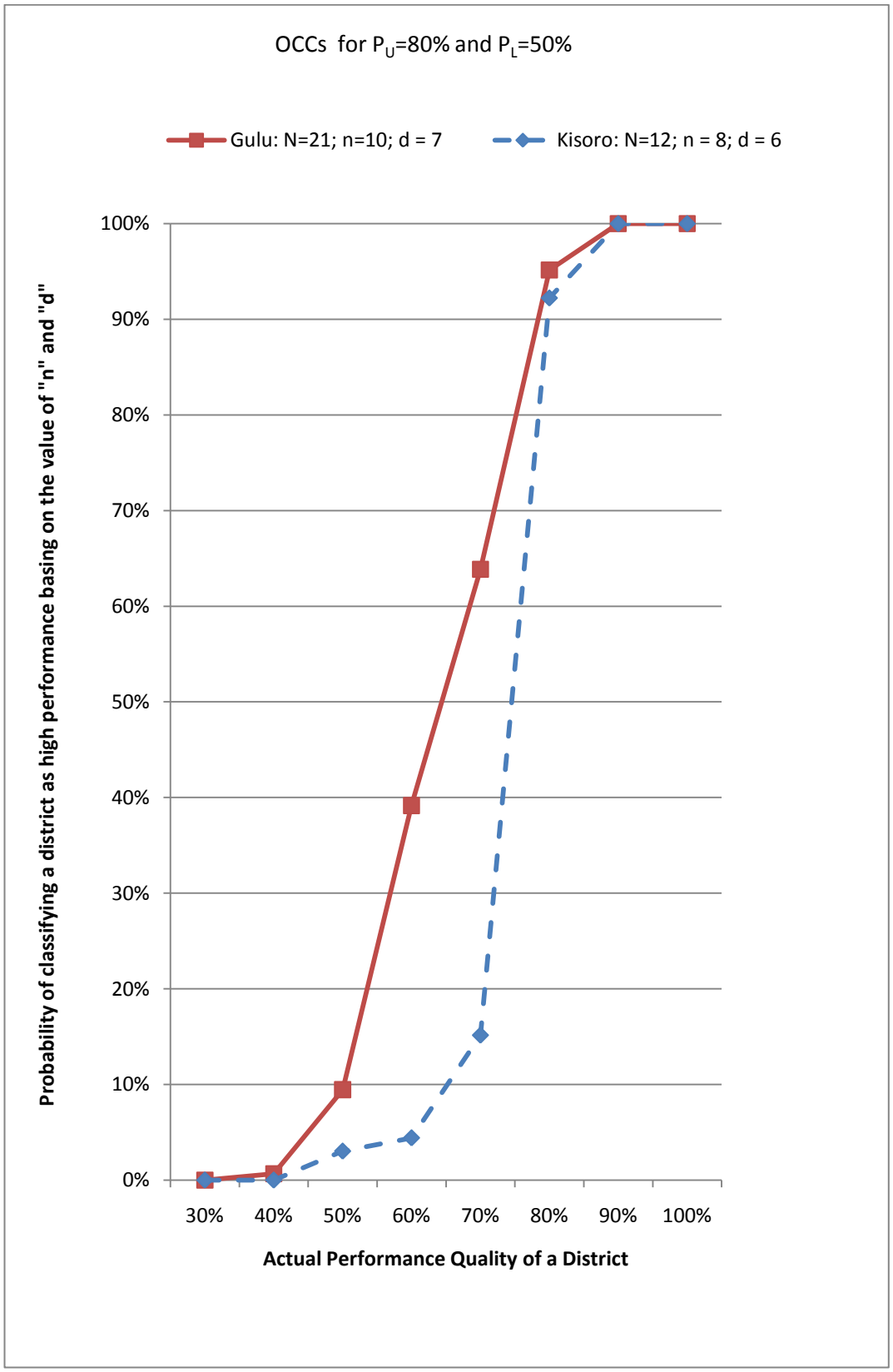
Decision rule (d) refers to the minimum number of HCs in a district required to meet the desired performance target (upper threshold) for the district to be classified as high performance. In Gulu district, for example, if ≥ 7 out of the sampled 10 HCs were found to meet the performance threshold for an indicator (e.g. patient assessment), the district as a whole was judged to be high performance with regard to this indicator. It was assumed there was no justification to not classify the district in the high category (which is that at least 80% of the 21 HCIIs in the district used the guideline as expected). This judgement is subject to risks of misclassification (α and β errors) of 0.049 and 0.095, respectively. Therefore, with n = 10 and d = 7, there is at least a 95.1% (1- α) probability of correctly classifying Gulu district as high performance with regard to a particular indicator if indeed ≥80% of all the HCs in the district met the performance benchmark for the specified indicator; and only a 9.5% probability of misclassifying it as high performance if only ≤ 50% of all the HCs in the district meet the performance threshold. A similar interpretation applies to Kisoro district with n = 8 and a decision rule of 6; and α and β errors of 0.078 and 0.030 respectively. LQAS method identifies the ends of the distribution. Therefore, the computed sample sizes (n)

and the decision rules (d) optimise the identification of the *best of the best* and *worst of the worst* performance districts with a small amount of error[9].

The Operating Characteristics Curve (OCC) in Figure 27 below shows the probability of classifying a district either as high or as low performance basing on the respective values of "n" and "d", and the true performance quality of the district against a given indicator. If the true performance quality of the district against an indicator falls between 50% and 80% thresholds, it is likely to be classified as high or as low performance depending on how close the performance is to the thresholds.

[9] In Public Health, we are more interested in identifying those who are at high risk—the worst of the worst

Figure 27: Operating Characteristic Curves for HFAs in Gulu and Kisoro districts for sample sizes of 10 and 8 with decision rules of 7 and 6 respectively



Selecting the HCs

Eligible HCs consisted of government-owned HCIIs where use of RDT-based guidelines had been introduced by the Uganda MOH. HCIIs with functional microscopes were excluded because they were in the process of being upgraded to HCIIIs. Eligible HCs were selected through stratified random sampling, with counties representing the strata. Random selection was undertaken by drawing names of the eligible HCs from a hat. The number selected from each county reflects the proportionate distribution of eligible HCIIs. In Kisoro district, the team did not observe the required number of patients (n = 6) in 2 of the HCs initially recruited into the study. The survey team replaced them by selecting other HCs which were nearest to them.

b) Sampling HWs and consultations

Sample size

In order to assess the HC, study assessed the performance of the HW who was most experienced in providing clinical care to a child with fever at the HC. LQAS was used to calculate the number of consultations (n) to observe in order to judge the performance of the individual HW, and to obtain the decision rule (d) for guiding the judgement. In this context, "d" refers to the minimum number of correct tasks the HW was expected to perform out of a set of consultations, in order for the HW to be classified as high performance. Because the number of children presenting with fever at HCIIs is presumed to be infinite, the study uses the binomial formula to calculate "n" and "d" [104, 106]. Calculation was based on four pre-set parameters:

i. Performance threshold (pU): the HW was expected to perform the selected clinical tasks according to the RDT-guidelines at least 95% of the time.
ii. Lower threshold (pL): the level of performance deemed seriously below the performance threshold that it was considered highly "unacceptable". This threshold was set at 50%. The lower threshold is set at 50% because we assume that the distribution of health worker performance is bimodal. Health workers either know the correct technique and use it, or they are ignorant of it and perform accordingly most of the time.
iii. The risk of judging a HC as a failure when in fact the observed HW achieved the performance benchmark (α error). The maximum permissible α-error was set at 0.10.

iv. The risk of judging a HC as adequately adhering to the guideline when in fact the observed HW performed at the pL threshold (β error). The maximum permissible β-error was set at 0.10.

Decision rule for HW performance

On the basis of these standards, the LQAS rule is to assess the HW performance for a set of tasks carried out on 6 patients (n = 6). For every 6 observations, no more than one error was permitted for any task or indicator. That is, the decision rule (d) was for the HW to perform a given task as expected at least 5 out of 6 times; otherwise the HC was judged as having failed to adhere adequately to the step or component of HC specified in the guideline. This judgement is subject to risks of misclassification (α and β errors) of 0.033 and 0.109 respectively. Therefore, this 6:5 design is 96.7% sensitive for identifying HWs that use recommended clinical norms 95% of the time, and 89.1% specific for HWs using the recommended norms ≤ 50% of the time. On the basis of this HW classification, the proportion of HCIIs in a district judged as adequately adhering to RDT guidelines was computed, and the result used to judge the performance of the district as a whole as explained in section 5.2.3 above.
This application of LQAS using binomials and the standards quoted above have been used in several surveys [104, 106, 127, 128] and were discussed with officials of the Ministry of Health and the districts prior to their application in Gulu and Kisoro districts.

c) Selecting patients

The study selected 6 consecutive consultations that were deemed eligible for observation of performance of an individual HW. Eligible patients consisted of children (aged 0 – 59 months old) attending outpatient clinics with uncomplicated fever, or a history of fever in the previous 24 hours.

Presence or a history of fever was a basis for inclusion into the study. A child was excluded from observation if s/he was showing danger signs, or signs of complications (Table 36 below). Use of RDT is not a pre-requisite in children presenting with severe illnesses [19, 157]. Fewer than 6 children could have been observed if children with severe illnesses were also included. However, we did not come across any child with symptoms and signs of severe illness.

Table 36: Danger signs of severe illness

- Convulsions or fits – now or within the past 2 days
- Not able to drink or breast feed
- Vomiting everything – not able to keep down food, fluid, or drugs
- Changes in mental state – patient is confused, very sleepy (lethargic or drowsy), or in coma (unconscious)
- Extreme weakness (prostration) – patient is unable to sit or stand without support
- Severe difficulty in breathing (respiratory distress)
- Severe anemia – pale palms, fingernails, eyelids
- Severe dehydration – coated tongue, sunken eyes, skin pinch

Source: Uganda ministry of health (user's manual)[157]

5.2.4 Data collection

a) Observing the HW

A trained research assistant familiar with the local language observed the HW seeing children with fever (assessment, use of RDTs, dispensing and counselling). The HW was requested to explain the purpose of this research to the caretakers of selected patients; and that an observer would be sitting in the consultation room to observe the consultation process. The HW invited the observer into the consultation room once consent was given by the caretaker. Each eligible patient whose caretaker consented to participating in the study was assigned a code number; which was recorded on a checklist used for the clinical observations, and on the instrument used to guide exit interviews. If, upon completion of the observation, the research team deemed that the observed child was ineligible (e.g. child was 5+ years; child had no history of fever), the relevant data were excluded and another observation carried out until 6 eligible consultations were observed[10].

Aspects of care assessed

Each HW was observed performing a total of 150 clinical tasks, 25 per child enrolled into the study. The tasks were drawn from 5 clinical procedures: (a) medical history (6 tasks), (b) clinical examination (6 tasks), (c) using RDT (11 tasks), (d) drug choice

[10] The research team took a brief history of the presenting complaints outside the consultation rooms as part of the exit survey.

relative to diagnosis (1 task) and (e) counselling/explaining treatment to patients (1 task). The specific tasks assessed are listed in Appendix 19.

Scales of measurement

The response to each item on the observation checklist was a "Yes", or "No" or "NA" (not applicable). A "yes" response means that the HW performed the expected task as recommended in the guideline; and a "no" response means s/he did not. A "yes" or "no" response was not applicable if, for example, a caretaker volunteered information about the presence of a symptom of interest (e.g. cough). In the final analysis, "Yes" and "NA" responses were combined into one category (a "pass" or "success"). Excluding "NA" responses would result in fewer than 6 responses per indicator. Basing on these categorical responses, the number of times each task was performed as expected (i.e. the number of passes or successes) was counted out of 6 observations. The 6:5 decision rule was then applied to judge the HW performance.

Judging HW performance across procedures

For a procedure consisting of multiple of tasks a HW had to perform all the individual tasks successfully in order to judge the entire procedure as having been performed adequately—i.e. the HW had to perform each of the constituent tasks correctly in at least 5 of 6 consultations in order to pass the entire set of tasks. For example, medical history taking consisted of 6 specific tasks (Appendix 19). If a HW asked about history of fever or history of cough in at least 5 of the 6 children, s/he was judged as having "passed" in performing those two tasks. On the other hand, if the HW asked about history of each of the other four symptoms in fewer than 5 children, s/ he was judged as having "failed" to perform each of the latter 4 tasks adequately. The HW in question was then judged as having failed in medical history taking as a whole. This procedure follows the LQAS principle for the second stage sampling of using d=5 as the decision rule. In short, in 6 observations at least 5 correct behaviours or responses are required. Table 39 below shows the number of successes per procedure, which represents the number of HWs that performed all the tasks included in a given procedure according to the specified standard.

b) Assessing adequacy of inputs

This section describes the protocol for assessing the availability of the following inputs which are essential for basic health facility functioning in the context of RDT-guidelines:

(a) essential drugs (b) RDT kits (c) Guidelines (d) equipment (e) staff. The assessment was carried out by observing the consultation rooms and stores, and by reviewing records of supplies and equipment. The decision rules for judging district performance (Table 35 above) were then applied to classify the districts as having either high or low performing health facilities for each of the specific indicator items assessed.

Availability of drugs

The indicator drugs chosen for this assessment were ACT for malaria, cotrimoxazole (CTX) for pneumonia or Acute Respiratory Infection (ARI), ORS for acute diarrhoea, ciprofloxacin (CIPRO) for dysentery, and Vitamin A (Vit. A). The study assessed availability of essential drugs on the day of the visit as well as over the 6 months preceding the survey[11]. A drug was considered available on the survey day if the study team observed at least one sealed basic unit (e.g. a tin or a packet) of the item. Items that were already open for use were not counted. In addition, the drug must have still been valid (not expired) to be considered available. ACTs which were distributed to the sampled HCs as part of the survey were not counted as part of the HC stock.

If the HC had several valid drugs and one or more non-valid drugs, then the HC failed this step. Furthermore, the study assessed whether the HC experienced stock-outs of any of the indicator drugs in the preceding 6 months—basing on the HC records and/or HW's report. If the HC experienced stock-out of any of the indicator drugs during the 6 months under review, then the HC failed this step.

Availability of RDT kits

Similarly, RDTs were considered available on the survey day if the study team observed at least one sealed box (10 tests) of the item, which must have still been valid (not expired). Items that were already open for use were not counted. RDT kits which were distributed to the sampled HCs as part of the survey were not counted as part of the HC stock.

Availability of guidelines

The study assessed the availability of the latest, nationally-mandated guidelines for the care of children with fever. Two types of fever-oriented guidelines were expected,

[11] In Uganda, supplies in government-owned HCs are normally replenished every 3 months; a 6-month time frame was considered to observe availability of inputs (especially of RDT kits) over a longer term. The degree to which competency deteriorates is related to the length of disruption of exposure to a new technology

namely, RDT-based guidelines and IMCI guidelines. Relevant material could be in the form of a booklet or a wall chart. A facility was considered to have all the mandated guidelines if it had both IMCI and RDT-related guidelines, in any form.

Availability of equipment

The study team checked the consultation rooms for availability of equipment which are essential in providing care to a child with fever. The assessment focused on 5 items: an infant scale, an adult scale, a timer (for counting respiratory rates and timing RDT process), a thermometer and an ORS cup or jar. Each item was considered available if it was observed by the study team and deemed to be functional.

Staffing

The cadres of staff typically recommended at a HCII are enrolled nurses (1), enrolled midwives (1), and nursing assistants (2) [182]. Staff availability was assessed by comparing the number of staff recruited in different positions relative to the official requirements for HCIIs: whether all the required positions for a particular cadre of staff had been filled. The study did not assess the extent of absenteeism. Where a HC had a HW that was more qualified than is officially recommended—e.g. a registered nurse, instead of an enrolled nurse, the relevant position was considered to have been filled.

c) Assessing adequacy of support activities

The study assessed HW exposure to in-service training and support supervision—two activities considered essential for introducing and scaling up the use of a new guideline. HWs who were observed providing care to the sick child on the day of the survey were interviewed about the dates of the most recent in-service trainings in fever management that they attended, and the content of the trainings. In addition, they were interviewed about exposure to support supervision in the preceding 3 months as well as the particular types of support offered during the visits. The decision rules for judging district performance (Table 35 above) were applied to classify the districts as having either high or low performing HCs with regard to the activities assessed.

d) Assessing adequacy of counseling

The study team interviewed care takers of enrolled patients as they left the clinic. The exit interview was conducted using a short structured questionnaire and checklist to

determine whether the caretaker was given advice on the child's diagnosis, the prescriptions, or the dosing. During the interview, the team further documented the details of the drugs prescribed to the child, including the dosing. This latter information was used to judge whether the diagnoses indicated by the HW matched the treatment prescribed.

5.2.5 Data collection tools

The survey used a tool adapted from the HFA tool for assessing the quality of care and health systems support for the Integrated Management of Childhood Illness (IMCI) policy. As already mentioned, the IMCI-HFA tool consists of 4 modules: (a) checklist for clinician observation of treatment of sick children, (b) questionnaire for exit interviews (c) checklist for assessment of infrastructure, equipment, drugs and supplies (d) questionnaire for health worker interview [170]. Because the focus of the research was on the quality of fever case management, the tool was modified by incorporating items which reflect the standard procedures described in the RDT-based guideline for managing fever [157]. Specifically, the adaptations involved

a) including steps in carrying out an RDT
b) including items for assessing RDT storage
c) excluding items for assessing inputs (drugs, supplies and infrastructure) that were deemed to have no direct effect on implementation of RDT-based guidelines (e.g. presence of vaccines, presence of a pit latrine)

5.2.6 Training of research assistants

In each district, data collection was preceded by a one-day training of a Research Assistant (RA). The RAs were experienced clinical officers with experience in survey methods. The RAs advised on clarity of questions, typographical errors, ease of translation of questions, and the reliability of the items. The training included demonstrations on the use of RDT kits and role play. In addition, the RAs participated in the final selection of HCs and scheduling of activities. Recommended changes were few and minor. The RAs carried out clinical observations and exit interviews as they were fluent in the local languages. I implemented the other modules of the HFA tool.

5.2.7 Field-testing of tools

The generic IMCI-HFA tools had been field-tested [170] and used before [92, 171, 172]. It was also used in Uganda (Gulu) in November 2010. However, the adapted, fever/malaria-specific version used in this study was pre-tested at one HC during the training of the RA in Gulu district in December 2010.

5.2.8 Data management and analysis

Data were independently entered into 2 separate data bases in MS Excel 2007 and checked for data entry errors using the COMPARE subroutine in Epi Info 3.5.2. Discrepancies were corrected by referring to the questionnaires. The analysis used SPSS (PASW statistics 18) to compute performance scores, and Excel 2007 to generate relevant tables and graphs.

5.2.9 Ethical considerations

The study protocol was reviewed and approved by the Ethics Committee in Liverpool School of Tropical Medicine and by the Uganda National Council for Science and Technology (UNCST). The UNCST wrote introductory letters to the respective District Authorities, which were used to secure verbal clearance from the latter. The research team explained the purpose of the research and the procedures to the health workers. The health workers did the same to the caretakers, in particular about the presence of the RAs in the consultation room. The team obtained consent from the health workers and, through the health workers, from caretakers before proceeding with observation and other aspects of the survey. The information and consent form can be found in Appendix 17

5.3 Results

5.3.1 Characteristics of the population

The surveys were conducted in Gulu (district in high malaria transmission zone) and Kisoro (district in low malaria transmission zone). The surveys were carried out in November and December of 2011.

On the day of the survey, nursing assistants and enrolled nurses and midwives constituted most (70%) of the providers of care to children with fever in Gulu district. The same cadres of HWs were the sole providers (100%) of care to children in Kisoro district (Table 37 below).

Table 37: Qualification of HWs who attended to the sick child in the surveyed health centres

	Gulu (n = 10)		Kisoro (n = 8)	
Qualification of attending clinician	Number	%	Number	%
Clinical officer	1	10.0%	0	0.0%
Registered nurse/midwife	2	20.0%	0	0.0%
Enrolled nurse/midwife	4	40.0%	3	37.5%
Nursing assistant	3	30.0%	5	62.5%

The study observed a total of 108 outpatient consultations from 18 HCs in 4 weeks: 60 consultations from 10 HCs in Gulu and 48 consultations from 8 HCs in Kisoro. The consultations consisted of children (0 – 59 months of age) presenting with uncomplicated fever. Table 38 shows the diagnoses made by the health workers from these consultations.

Table 38: Health workers' diagnoses by district

DIAGNOSES	Gulu district (High transmission) (60 consultations)	Kisoro district (Low transmission) (48 consultations)
Acute respiratory infections (ARI) alone	46.7%	27.1%
Malaria alone	18.3%	14.6%
Malaria and other diagnosis	26.7%	29.2%
Other conditions (excluding malaria & ARI)	8.3%	29.2%

In Gulu (high transmission), the vast majority (91.7%) of the consultations was classified either as Acute Respiratory Infections (ARI) alone (46.7%) or as malaria (45%)—either alone or as a co-diagnosis. All the diagnoses specified as ARI and/or

malaria were based on the use of RDT. Other conditions were considered where RDTs were not used.

In Kisoro (low transmission), most (43.8%) of the consultations were classified as malaria (alone or as a co-diagnosis). A substantial number (29.2%) of patients were classified into conditions other than malaria or ARI. Most diagnoses in Kisoro (low transmission) were not based on RDT results. Only 29.2% of the diagnoses in the district were accompanied by use of RDTs.

5.3.2 Health worker performance

a) Performance by procedures

In both districts, none of the HWs performed all the 5 procedures satisfactorily. Only two procedures were performed to satisfactory standards: treatment prescribed (Gulu, high transmission) and Counselling (Kisoro, low transmission). Treatment matched the diagnosis indicated by the HW in at least 5 of 6 consultations at 7 HCs in high transmission area (Gulu). Thus the high transmission district (Gulu) was judged as high performance with regard to appropriateness of treatment prescribed (d = 7). In low transmission area (Kisoro), counselling was provided as expected—to at least 5 of 6 care takers per HC—at 6 HCs. Therefore, the low transmission district (Kisoro) was judged as high performance with regard to care taker counselling (d = 6). Both treatment and counselling are one-task procedures. Performance of multi-task procedures—namely medical history, clinical examination and use of RDT—was inadequate in both districts; i.e. none of the HWs in either district performed all the essential components of each of these procedures to satisfactory standards. Notably, at only 2 of the HCs in high transmission area (Gulu) did the HWs follow all the required steps in using a RDT in at least 5 of 6 consultations. In low transmission area (Kisoro), none of the HWs used RDT satisfactorily in at least 5 of 6 consultations.

Table 39: Performance of HWs by healthcare procedures

		GULU (High transmission) (n=10; d = 7)		**KISORO** (Low transmission) (n=8; d = 6)	
Procedure	**Description**	*Success	**District performance	*Success	**District performance
Medical history	HW asks about all essential symptoms	0	Low	0	Low
Clinical examination	HW performs all essential clinical examination tasks	0	Low	0	Low
Use of RDT	HW follows all required steps in using RDT	2	Low	0	Low
Treatment	Treatment matches diagnosis	7	High	4	Low
Counselling	HW explains treatment to participants	6	Low	6	High
All procedures	HW performs all observed clinical tasks as required	0	Low	0	Low

** Number of health workers that performed all the tasks included in a given procedure according to the specified standard*

***High = satisfactory or adequate performance; Low = unsatisfactory or inadequate performance*

n = sample size (HCs); d = decision rule

b) Performance by specific tasks

This section describes HWs' performance for each of 25 clinical tasks observed. This analysis identifies the most problematic steps in managing fever in children. The results are displayed in Table 40 below.

Medical history: within the category of medical history, the tasks with low successes in both districts were: asking about feeding, asking about vomiting and asking about convulsions. That is, HWs rarely asked about symptoms which are indicative of severe forms of illnesses in children.

Clinical examination: None of the HWs observed met the performance threshold for any of the clinical examination tasks. Hence, adherence to all clinical examinations tasks was low in both districts

RDT performance: In high transmission area (Gulu), utilisation of RDT was high. Most RDT-related tasks were performed satisfactorily. The most problematic steps were (a) checking the expiry date; (b) writing the patient's name on the cassette; (c) allowing fingers to dry before pricking; and (d) waiting at least 15 minutes after buffer (HWs often read results within 2 minutes after adding buffer).

In low transmission area (Kisoro) utilisation of RDT was low. Performance of all the essential tasks required in carrying out a RDT was unsatisfactory, mainly due to errors of omission. HWs rarely considered malaria as a possibility and rarely used RDT if the patient presented with features of ARI; except if such a patient looked ill or had high fever at the point of consultation.

Drug choice: Treatment-diagnosis match was high in high transmission area (Gulu). It was low in low transmission area (Kisoro), partly because the HWs did not prescribe some of the medications required for the diagnoses they made, and because of prescribing ACT presumptively or to test negative patients.

Counselling: The number of patients counselled about the treatment prescribed was inadequate in high transmission area (Gulu). Counselling was high in low transmission area (Kisoro).

Table 40: Performance of HWs by specific clinical tasks

		GULU (High transmission) (n=10; d = 7)		KISORO (Low transmission) (n=8; d = 6)	
Procedure	**Tasks**	*Success	**District performance	*Success	**District performance
Medical history HW asks about:	Fever	10	High	8	High
	Cough or difficulty in breathing	8	High	4	Low
	Diarrhoea	2	Low	3	Low
	Feeding	1	Low	0	Low
	Vomiting	1	Low	0	Low
	Convulsions	0	Low	0	Low
Clinical examination HW checks for:	Vaccination status	0	Low	0	Low
	Dehydration	0	Low	0	Low
	Respiration rate	0	Low	0	Low
	Conjunctiva: pallor and jaundice	0	Low	0	Low
	Throat for redness	0	Low	0	Low
	Ear for discharge or sores	0	Low	0	Low
Use of RDT HW:	Requests RDT	9	High	1	Low
	Checks expiry date	6	Low	0	Low
	Puts on gloves	8	High	1	Low
	Writes Patent's name on cassette	5	Low	1	Low
	Cleans patient's finger	9	High	1	Low
	Allows finger to dry before prick	3	Low	0	Low
	Uses pipette to collect blood	9	High	1	Low
	Puts blood into position A	9	High	1	Low
	Puts buffer in position B	9	High	1	Low
	Waits at least 15 min after buffer	4	Low	0	Low
	Communicates RDT results	8	High	1	Low
Treatment	Treatment matches diagnosis	7	High	4	Low
Counselling	Health worker explains treatment	6	Low	6	High

** Success: number of health workers that met the performance threshold with regard to the specified task*

***High = satisfactory or adequate performance; Low = unsatisfactory or inadequate performance*

n = sample size (HCs); d = decision rule

5.3.3 Adequacy of inputs

a) Availability of essential drugs

This section describes the adequacy of essential drugs in terms of their availability in the health facility on the day of the visit, and during the 6 months preceding the survey.

Number of facilities with essential drugs in stock on survey day

Table 41 shows the percent of HCs with the drugs in question on the day of the survey.

Table 41: Performance of districts based on availability of essential drugs on survey day

	GULU (High transmission) (n = 10; d = 7)		KISORO (Low transmission) (n = 8; d = 6)	
Essential drugs	*Success	**District performance	*Success	**District performance
ACT for malaria	10	High	7	High
Cotrimoxazole for pneumonia	6	Low	4	Low
ORS for acute diarrhoea	7	High	8	High
Ciprofloxacin for dysentery	0	Low	1	Low
Vitamin A	9	High	8	High
All essential drugs	0	Low	1	Low

** Success: number of HCs with the listed drugs in stock on day of visit*

***High = satisfactory or adequate performance; Low = unsatisfactory or inadequate performance*

n = sample size (HCs); d = decision rule

Neither Gulu nor Kisoro had adequate amounts of all the essential drugs in stock on the day of the survey. In both districts, the most problematic drugs were cotrimoxazole for treating ARI and CIPRO for treating dysentery. Otherwise ACT, Vitamin A and ORS were available in adequate amount on the day of the survey.

Availability of essential drugs over the preceding 6 months

Table 42 below shows the number of HCs that did not experience any stock-out of the listed essential drugs. None of the districts had adequate stocks of all the essential drugs throughout the 6 months preceding the survey. In high transmission area (Gulu), the stock of most of the essential drugs was low during the 6 months period under review; except for Vitamin A. In low transmission area (Kisoro), the stocks of ACT, ORS and Vitamin A were adequate throughout the 6 months period prior to the survey; only CTX and CIPRO were in short supply during this period.

Table 42: Performance of districts based on availability of drugs during 6 months prior to survey

Essential drugs	GULU (High transmission) (n = 10; d = 7) *Successes	**District performance	KISORO (Low transmission) (n = 8; d = 6) *Successes	**District performance
ACT for malaria	6	Low	8	High
Cotrimoxazole for pneumonia	5	Low	4	Low
ORS for acute diarrhoea	5	Low	6	High
Ciprofloxacin for dysentery	3	Low	0	Low
Vitamin A	7	High	8	High
All essential drugs	0	Low	0	Low

** Success: number of HCs that did not experience stock-outs of the specified item*
***High = satisfactory or adequate performance; Low = unsatisfactory or inadequate performance*

n = sample size (HCs); d = decision rule

b) Availability of RDT

Neither of the districts had adequate stock of RDT. In both districts, RDT kits had been out of stock for an average of 8 months, and HWs had reverted to presumptive treatment. RDT kits which were supplied as part of the survey was not counted in the HC stocks.

c) Availability of guidelines

The study assessed the availability of the latest, nationally-mandated guidelines for the care of children with fever. A facility was considered to have all the mandated guidelines if it had both IMCI and RDT-related guidelines. Table 43 below summarises the number of HCs where the 2 types of guidelines were available (in any forms) and were easily accessible.

Neither of the 2 districts had the required guidelines in sufficient number. None of the job aids assessed was available in sufficient number in high transmission area (Gulu). Wall charts—explaining how to carry out a RDT procedure or how to treat a case with positive or negative results—were available in sufficient number in low transmission area (Kisoro). Otherwise, there was shortage of other types of key job aids in low transmission area as well.

Table 43: Performance of districts based on availability of mandated guidelines on survey day

Guidelines or wall charts	GULU (High transmission) (n = 10; d = 7) *Successes	**District performance	KSORO (Low transmission) (n = 8; d = 6)) *Successes	**District performance
Wall charts***	5	Low	6	High
IMCI algorithm (wall chart)	6	Low	3	Low
RDT/ACT user manual	2	Low	1	Low
IMCI guidelines	6	Low	1	Low
Both IMCI and RDT-based job aids	6	Low	2	Low

** Success: number of HCs with relevant guidelines*
***High = satisfactory or adequate performance; Low = unsatisfactory or inadequate performance*
****guide on how to carry out a RDT procedure/on treating RDT positive and negative patients*

n = sample size (HCs); d = decision rule

d) Availability of equipment

Assessment of the availability of essential equipment was based on 5 items required in the management of a child with fever, namely infant scale, adult scale, timer (for counting respiratory rates and timing RDT process), thermometer and an ORS cup/jar. Table 44 displays the results.

Table 44: Performance of districts based on availability of equipment on survey day

Essential equipment	GULU (High transmission) (n = 10; d = 7) *Successes	**District performance	KISORO (Low transmission) (n = 8; d = 6) *Successes	**District performance
Infant scale	10	High	8	High
Adult scale (for older children)	5	Low	5	Low
Timer	7	High	3	Low
Thermometer	7	High	2	Low
ORS cup/jar	1	Low	1	Low
All essential equipment	1	Low	0	Low

**Success = number of HCs with the relevant equipment*
***High = satisfactory or adequate performance; Low = unsatisfactory or inadequate performance*

n = sample size (HCs); d = decision rule

Neither of the 2 districts had sufficient stock of all essential equipment at the time of the survey. In Gulu, the problematic items were adult scales and ORS cup/jar. In

Kisoro, all the listed pieces of equipment were in short supply, except infant scales which were in sufficient number.

e) Availability of sanctioned staff

Staff availability was assessed by comparing the official staffing requirements for HCIIs versus the positions filled for different cadres. The results are summarised in Table 45 below. Gulu district (high transmission) had sufficient number of professional staff while Kisoro district (low transmission) had sufficient number of both professional and paraprofessional staff. However, none of the HCs in either district had sufficient number of all the required cadres of staff.

Table 45: Performance of districts based on availability of required staff

	GULU (High transmission) (n = 10; = 7)		KISORO (Low transmission) (n = 8; d = 6)	
Types of staff	*Successes	**District performance	*Successes	**District performance
Professional staff (enrolled nurse/midwife)	7	High	6	High
Paraprofessional staff (nursing assistants)	5	Low	6	High
All recommended staff	3	Low	5	Low

** Success: number of HCs with required number of staff as per the staffing norm of 2009 (Human resource for health: Uganda [182])*
***High = satisfactory or adequate performance; Low = unsatisfactory or inadequate performance*

n = sample size (HCs); d = decision rule

5.3.4 Adequacy of support activities

a) In-service training in fever management

The HWs who provided care to the sick child on the day of the survey were interviewed about the dates of the most recent in-service trainings in fever management that they attended, and the content of the trainings. Table 46 below describes their responses.

Training in fever case management was adequate in high transmission area (Gulu). Notably, exposure to most aspects of the RDT-based guideline was found to be

adequate. The aspects that were reported to have not been adequately covered included (a) managing a case with a negative RDT test, and (b) managing a case with ARI.

Table 46: Performance of districts based on training in fever case management

	GULU (High transmission) (n = 10; d = 7)		KISORO (Low transmission) (n = 8; de = 6)	
	*Success	**District performance	*Success	**District performance
Most recent training attended by provider				
HWs receiving training in past 1-3 years ago*	7	High	4	Low
Content of training				
Evaluating patient with fever	7	High	4	Low
Selecting patient for RDT	7	High	4	Low
Performing RDT/reading RDT result	7	High	4	Low
Managing patient with positive RDT result	7	High	4	Low
Managing patient with negative RDT result	3	Low	4	Low
Recognising/referring severe illness	7	High	4	Low
Patient education/counselling	7	High	4	Low
RDT storage/monitoring	7	High	4	Low
Treatment with ACT	7	High	4	Low
Management of pneumonia/ARI	6	Low	1	Low

**Success: number of HWs trained. In Gulu, 6 HWs trained in previous 12 months; in Kisoro, all 4 HWs trained between 1to 2 years previously*
***High = satisfactory or adequate performance; Low = unsatisfactory or inadequate performance*
n = sample size (HCs); d = decision rule

Training in fever case management was low in low transmission area (Kisoro). Consequently, exposure to all the relevant aspects of the RDT-based guideline was inadequate.

b) Supervision

The study assessed both the frequency of support supervision offered to the interviewed HW, and the particular types of support offered during such visits. The results are summarised in Table 47.

Exposure to support supervision during the 3 months preceding the survey was high in both districts. In Gulu, exposure to most of the expected support activities was adequate, notably observing and giving feedback on the HW's work. The problematic support activities were drug delivery and giving updates on policies. In Kisoro, most of

the expected support activities were not adequately provided, except for checking of records and observing the HW's work. Feedback on the HW's work was inadequate.

Table 47: Performance of districts based on number and quality of supervision

	GULU (High transmission) (n = 10; d = 7)		KISORO (Low transmission) (n = 8; d = 6)	
Most recent visits	*Success	**District performance	*Success	**District performance
Supervision in last 3 months	10	High	8	High
Types of support offered				
Delivered supplies	6	Low	1	Low
Checked records	8	High	8	High
Observed provider's work	8	High	8	High
Gave feedback	8	High	5	Low
Gave update	6	Low	5	Low
Discussed problems	7	High	4	Low
Checked drugs supply	9	High	4	Low

**Success: number of HWs who were supervised in the last 3 months/who received the specified support*
***High = satisfactory or adequate performance; Low = unsatisfactory or inadequate performance*

n = sample size (HCs); d = decision rule

5.3.5 Summary of findings

Table 48 shows a score card for the two districts for the 4 inputs items and the 2 process items included in the assessment.

HW performance

None of the HWs met the performance threshold for all the tasks observed. Therefore, both districts failed to adequately adhere to the recommendations of the RDT-based guidelines. Adherence was high in Gulu for prescribing the correct medication for the diagnosed illness and in Kisoro for client-counselling —all of which were single-task procedures. Medical history taking and clinical examination were seriously incomplete in both districts. Notably, HWs rarely asked about symptoms that are indicative of the severe forms of illness in children with fever. RDT was performed in 91.7% of consultations in Gulu—a high malaria transmission setting; compared to 29.2% of consultations in Kisoro—a low malaria transmission setting.

a) Adequacy of essential inputs and support activities

None of the districts had adequately stocked all the essential drugs either on the day of the survey or during the preceding 6 months. The most problematic drug items were cotrimoxazole (for treating pneumonia or ARI) and CIRPO (dysentery)—both on the survey day and during the 6 months prior. ACT was notably available in adequate amount in both districts on the survey day; and during the preceding 6 months in Kisoro. The stock of RDT was unacceptably low in both districts, both on the day of the visit and throughout the 6 months preceding the survey. In addition, there were insufficient numbers of relevant guidelines as well as essential equipment in both districts. Although none of the districts had adequate number of all the recommended staff, both districts had adequate number of professional staff. In-service training on fever case management (in the context of RDT) was high in high transmission area (Gulu): the topic which was addressed insufficiently was the management of a patient with a negative RDT result. In low transmission area (Kisoro), training was generally inadequate, and all topics were insufficiently covered. Support supervision was adequate in both districts.

Table 48: Overall performance of the two districts assessed

Domain	Input/ activity	Indicators	**Gulu** (n = 10) (d = 7)*	**Kisoro** (n = 8) (d= 6)*
HW performance	Medical history	HCs where HW ask about all essential symptoms	0 Low	0 Low
	Clinical examination	HCs where all essential clinical examination tasks are assessed	0 Low	0 Low
	Use of RDT	HCs where HW follow all required steps in using RDT	2 Low	0 Low
	Treatment	HCs where all treatment prescribed are appropriate to the diagnosis	7 High	4 Low
	Counselling	HCs where HW explains treatment to all care takers	6 Low	6 High
	All procedures	HCs where HW performs all observed clinical tasks as required	0 Low	0 Low
Inputs	Drugs	HCs with all (5) first line drug items for childhood fever in the surveyed HCs on the day of the survey (ORS, CTX, CIPRO, ACT, VIT A)	0 Low	1 Low
		HCs with no stock-outs of the essential drug items in the preceding 6 months	0 Low	0 Low
	RDT	HCs with at least one basic unit (sealed pack) of RDT kits available on day of visit	0 Low	0 Low
	Guidelines	HCs with all nationally-mandated guidelines for managing a child with fever, being available and accessible on the day of the survey (IMCI guideline/wall chart (algorithm) and RDT/ACT user manual or wall chart guide on performing RDT + wall chart on treatment according to RDT results)	6 Low	2 Low
	Equipment	HCs with all essential equipment required to support the management of a child with fever on day of survey (infant scale, adult scale, timer, thermometer, ORS cup/jar)	1 Low	0 Low
	Staffing	HCs with all the sanctioned clinical staff as per the day of the survey	3 Low	5 Low
Support activities (Processes)	Training	HCs in which interviewed HWs reported receiving training on guidelines for fever in the preceding 1-3 years	7 High	4 Low
	Supervision	HCs in which interviewed HWs reported receiving support supervision in the preceding 3 months	10 High	8 High

**d: decision rule*

5.4 Discussion

This chapter evaluates the quality of implementation of RDT-supported policy in a district as a whole. It assesses the quality of delivery of all components of the guideline as well as adequacy of key health systems inputs and support activities. Assessment of HW adherence is based on observation of consultations of children (< 5years) presenting with fever. The HW was observed performing 4 clinical procedures (a) clinical assessment (medical history and clinical examination), (b) use of RDT, (c) treatment prescribed, and (d) counselling of care takers about therapy. The survey compares HW adherence from high and low malaria transmission settings. It identifies the problematic elements of the health care process which may need strengthening. It also assesses the adequacy of relevant inputs and support activities essential for effective implementation of guidelines.

Since use of RDT-based guidelines is intended to improve patient care, particularly of non-malarial fevers [19, 70-72], the quality of the diagnoses made and the appropriateness of treatment prescribed by the HWs are priority problem areas in the assessment of HW performance. The quality of diagnosis and type of treatment prescribed also impact on healthcare cost. Accordingly, the following tasks were considered critical in evaluating HW adherence to RDT guideline in Uganda (a) assessment (medical history and clinical examination), which includes all tasks except checking for vaccination status; (b) use of RDT, which includes requesting and using RDT in all fever patients, use of a loop or pipette to collect the blood, putting blood and buffer in the appropriate wells and waiting for at least 15 minutes after buffer before reading the test results; and (c) the type of treatment prescribed. Counselling of care takers is important in ensuring effective treatment. It does not directly influence the quality of diagnosis and the type treatment prescribed. Therefore, it was not considered a priority problem area.

5.4.1 Health worker adherence

The results show that neither of the districts implemented the guideline sufficiently. Adherence was particularly low in the diagnostic steps: medical history taking and clinical examination. Notably, HWs very rarely asked about symptoms that are indicative of the severe forms of illness in children with fever. In addition, and

surprisingly, use of RDT was remarkably lower in Kisoro—a low malaria transmission setting—than in Gulu—a high malaria transmission setting.

Previous evaluations of implementation of IMCI guidelines—including one from Uganda—have also shown low performance scores, with patient counselling being the most problematic procedure and treatment prescribing showing the best scores[90, 92, 93, 113]. Performance of assessment tasks has been found to vary widely across studies, from 27.3% in Benin [115]to 86.9% in Morocco [90], averaging at 49.8% (95% CI: 49.3% to 50.3%). In this survey, the pooled scores for assessment tasks (medical history and clinical examination) were 16% in Kisoro and 25% in Gulu. The findings in this survey show that performance of assessment tasks in the districts assessed appears to be on the poorer side compared to the performance reported in several surveys of IMCI guidelines [90, 92, 93, 113]. It is acknowledged that meaningful comparison of the findings of this survey with previous studies is hampered by variability in context; number of tasks assessed and design effects, which might be larger than anticipated.

It appears that, in managing a case of fever, HWs tend to focus more on use of RDTs and/or correct prescribing of antimalarials relative to RDT results than, on carrying out complete clinical assessment and other components of the guideline. This might be because of the way RDT-based guidelines have been portrayed during trainings, which tend to emphasize more of the former than the latter, as was noted with the trials synthesized in chapter 3 [34-37]. Several other training programmes have also had similar unbalanced emphasis on implementation of different components of RDT-based guidelines [32, 120, 121, 151]. This survey has also shown that trainings did not adequately address management of patients with negative RDT results, even in the district where in-service training was found to be adequate.

By emphasising use of RDTs and prescribing of anti-malarials more than other components of the guideline, HWs might have a mindset that RDT-based guidelines are a tool for preserving ACTs or antimalarials. Relative to presumptive treatment, use of RDTs probably does not improve care in parasite positive cases of fever since some of them are missed [138]. The main clinical advantage of using RDTs in fever case management is in aiding the identification of parasite negative cases in whom alternative diagnosis can be sought [19, 70-72]. This depends on complete assessment of patients

[138, 155, 157, 166]. Guidelines and training programmes need to be very clear about this objective.

This survey found a stark difference between the two districts in the use of RDTs. Use of RDTs was high in high transmission area and low in low transmission area. In high malaria transmission area, HWs used RDT in 91.7% of the consultations. In low transmission area, HWs used RDT in only 29.2% of the consultations. Use of RDTs in fever case management is potentially cost-saving and cost-effective in low to moderate malaria transmission settings, as shown by the results of the analysis of trial data from Uganda in chapter 4, and by previous economic models[12] [12, 13, 33, 72, 140, 144]. Both the analyses in chapter 4 and previous economic models show that use of RDTs in high malaria transmission settings is unlikely to be cost-saving, or cost-effective, even if HW adherence to the guideline is perfect. Therefore, the pattern of utilisation of RDTs observed in this survey—showing low utilisation of RDT in low transmission district and high utilisation of RDTs in a high transmission district—is contrary to expectations. A similar pattern of use of RDTs in fever case management was observed in Kenya, and was found to be associated with substantial increase in healthcare cost in a low transmission setting [33]. If this pattern is replicated more widely, then RDT-based treatment policies for fever may lose their economic advantage over presumptive treatment, even in low and moderate transmission areas, in Uganda. The substandard quality of assessment exacerbates the overall quality of care in both settings.

The pooled performance scores[13] in this survey was 52% in high transmission area (Gulu) and 21% in low transmission area (Kisoro)(Appendix 20). These scores are much lower than those obtained from correct use of RDTs and prescribing of antimalarials as the sole indicators of the quality of implementation of RDT guidelines [35-37, 120, 121]. Assessments based on prescribing alone tend to give higher adherence scores [90, 91, 109, 110, 112, 146, 150] than those based on multiple procedures/indicators [33, 92, 93, 113, 115, 117]. HWs tend to prescribe treatment according to the diagnoses they indicate. Hence it is logical for treatment scores to be high. Therefore, the findings of this survey give a more accurate picture of the quality of implementation of the RDT-

[12] Literature review: section 2.8.7
[13] Percent of all the tasks that were performed according to guideline

based guideline in routine practice in Uganda. It indicates that the problem of non-adherence is much deeper than is normally reported.

5.4.2 Adequacy of inputs and support activities

The findings of this evaluation suggest that the health system is not adequately supportive of the routine use of the RDT-based guideline. Notably, RDT kits were found to have been out of stock for 8 to 9 months in both districts and HWs had reverted to presumptive treatment at the time of the survey[14].

Inadequacy of inputs may not explain fully the poor HW performance in both districts. For example, although both districts had insufficient number of guidelines, the effect of accessibility to guidelines on HW performance is generally equivocal [93, 112, 114, 146, 150][15]. Availability of essential drugs does not appear to be the reason for poor HW performance either. Prescribing of treatment was appropriate in high transmission area (Gulu). Inappropriate treatment in low transmission area (Kisoro) was mainly due to prescribing of anti-malarials presumptively and due to polypharmacy (prescribing more drugs than was needed for the diagnosis indicated by the HW).

Shortage of RDT over a prolonged period of time may have had the biggest impact on (low) guideline implementation. Human capital theory underlines the importance of consistency of use of a new technology in order to maintain proficiency over time[16] [28, 31, 41, 42, 46], especially if its use requires the acquisition of new skills [41]. An LQAS assessment of family planning services in Kenya found that service providers who continued to use their skills had better skills-retention level than those who were no longer offering the same services at the time of the survey [106]. Therefore, the prolonged disruption in the use of RDTs could have resulted in loss of interest and loss of skills in its use [60, 61, 106, 183, 184]. However, this factor per se cannot explain the difference in performance quality between the two districts since the problem of shortage of RDTs was general and appears to have affected both districts equally.

Exposure to in-service training was low in low transmission area (Kisoro); and this may have contributed further to poor performance in the district—although the evidence on

[14] RDT kits which were supplied as part of the survey was not counted in the HC stocks

[15] See Literature review, section 2.10.2

[16] See Literature Review, section 2.3.2

the effect of in-service training on adequacy of adherence to guidelines is considerably variable [90, 91, 93, 112, 117, 146]. Further, lack of refresher trainings could have affected performance quality in both districts. The positive effect of in-service training on adherence to guidelines is generally modest [185] and short-lived [62]; and needs to be topped-up and reinforced by regular refresher trainings and constant exposure to the innovation.

In a nutshell, shortage of job aids, lack of refresher trainings and prolonged shortage of RDTs could all explain the poor performance quality in both districts. Inadequate exposure to initial in-service training may explain the poorer performance quality in low transmission area (Kisoro).

5.4.3 Usefulness of the findings

The survey presents performance scores based on a comprehensive and simultaneous assessment of all aspects of the healthcare process as stipulated in the RDT-guideline. Therefore, it presents a more comprehensive measurement of adherence to RDT-guideline [109, 113]. Accordingly, it portrays the quality of the entire process of health care for patients with fever in the context of the new guideline [109, 113].

The results of this survey helps to identify elements of the healthcare process with a particular need for improvement or more emphasis [113]. In both Gulu and Kisoro districts, improvements are required in the clinical assessment component of the guideline (medical history and clinical examination) in order to improve the quality of care, especially in non-malarial fevers. In Kisoro, particular emphasis is required on use of RDTs to support diagnosis in fever patients and on the need to use the results to guide prescribing of medications. In Gulu, particular improvement is needed in counselling of care takers. These improvements can be brought about through refresher trainings and/or support supervision, and by making job aids available and accessible to HWs. It is imperative for health authorties to main an adequate level of RDT kits at the facilities at all times for the suggested improvements to take place and for the new guideline to be implemented to acceptable quality.

To my knowledge, this is the first study in Africa that provides a comprehensive picture of the quality of implementation of the RDT-based guideline, alongside a systematic

assessment of the systems readiness in supporting the routine implementation of the guideline.

Furthermore, this study is unique in the sense that it uses LQAS method which has the advantage of identifying local variation in performance indicators (e.g. facility-specific performance of clinical tasks), while collecting data for regional and national assessments [106]. An example of facility-specific performance is shown in Appendix 19 for HW performance indicators. Therefore, interventions aimed at improving HW performance can be targeted at both district and facility levels. For indicators measured at district level only (e.g. adequacy of health systems inputs), interventions can not be targeted at facility levels.

The LQAS shows how the 6:5 observation rule can give a quick portrait of the quality of different aspects of care; and how the decision rule can be used alongside the checklists for HFA as a supervisory tool [106].

5.4.4 Study limitations

This study focuses on children only and could be biased. However, there is no evidence from this study that the age of the patients was a factor in deciding whether or not to perform an RDT. When an RDT was used, HWs adhered strictly to the results. Because it focuses on children, the data generated is of limited value in modelling the cost implications of using RDT-based guidelines in the general population.

The sample size in the LQAS method is too small to measure coverage (point estimates) at a facility level, but could be combined to measure coverage at the district level [106, 186]. In this analysis, the pooled number of observations per district was still too small to give accurate point estimates of district-level adherence scores. As such these (pooled point estimates) may not be useful in economic evaluation models.

The study shows high adherence scores for the quality of treatment prescribed—measured as the match between diagnosis and drug choice. However, LQAS-based methods only describe whether or not a patient was classified into a diagnostic category. The quality of diagnoses upon which treatment is based may be conditional on the quality of prior clinical assessments and RDT test [113]. Additionally, an important element of the quality of treatment in a child includes the dosing regimens of the medications prescribed—which are age and weight dependent, particularly the dosing

regimen of ACTs [17, 157, 187]. No "gold standard" assessments of the patients were performed to judge the validity of the diagnoses made by the health workers; or to assess whether the dosages of drugs prescribed were appropriate for weight and age. Therefore, it is probable that the amount of errors associated with treatment choices was higher than is reported by this study.

Health providers are known to improve their practices once they are aware that they are being observed or their performance is being assessed (Hawthorne effect) [188, 189]. This implies that the level of HWs performance observed in this study may be better than it is in normal everyday practice. In spite of this, HWs performance was generally low in both districts, implying that HWs might execute even fewer of the required tasks when they are not being observed. Further, because the survey uses the most experienced HW to represent a HC, the performance reported in this survey may be interpreted to be better than usual. In reality we observed the same staff that would normally carry out the same tasks in everyday practice. Therefore, the risk of bias associated with observing the "most experienced" HW is considered to be negligible. In any case, the performance of most tasks was low in both districts.

Because of language barrier, the RAs carried out clinical observations and exit interviews. As such there is a risk of observer error. The performance of the RAs during training suggests that the level of such a risk is negligible.

5.5 Conclusions

The results of this survey show that the quality of implementation of the RDT-based guideline through routine health services was inadequate in both high and low malaria transmission districts. Notably, implementation was much poorer in the low transmission district where the use of RDT is expected to be of a greater economic potential. Basing on the pattern of adherence, use of the RDT-based guideline may not be economically beneficial in routine clinical settings in Uganda.

Several areas of the guideline require improvement. Clinical assessment of patients was seriously incomplete in both districts, and needs to be emphasized through further interventions. Additional areas for improvement include counselling (Gulu-high transmission), use of RDT (Kisoro-low transmission) and prescribing (Kisoro-low transmission). Improvements can be brought about through refresher trainings, support supervision, provision of job aids and regular supply of RDT kits. These interventions could be undertaken prior or during dissemination of the guideline for wide-scale implementation. Without these improvements, use of RDT-based guidelines may not result in better management of fevers relative to presumptive treatment.

It is also evident that the health system is not adequately supportive of the routine use of the guideline and is in need of strengthening. Apart from ACTs, ORS and Vit A, most of the essential inputs required for the management of the sick child were lacking on the day of the visit. Most significantly, there was a prolonged disruption in the supply of RDTs in both districts. Several HCs were lacking necessary job aids, especially in Kisoro district. A number of HWs from both districts had not attended in-service training in the use of the guideline; whilst there was an urgent need for refresher training.

Chapter 6

General Discussion and Conclusions

Chapter 6 General discussion and conclusions

6.1 Overview

This thesis aimed to establish if implementing an RDT-based guideline instead of presumptive treatment of fever leads to better quality of patient care relative to presumptive treatment when delivered using carefully controlled studies and through routine clinical practice. In particular, the thesis

a) reviews evidence from RCTs to determine if RDT-based policies for fever lead to better quality of patient care than presumptive treatment under optimal conditions

b) analyses data from a pragmatic trial to determine if RDT-based policies for fever lead to better quality of patient care than presumptive treatment in routine practice

c) assesses the quality of actual practice when an RDT-based guideline is rolled out to health services in a district as a whole

The thesis consists of 3 studies which reflect the sequence of evaluation of a new healthcare technology, from assessment of efficacy, through assessment of effectiveness to assessment of current practice (Table 49). A Cochrane review of evidence from randomised trials in Africa is used to evaluate the effect of RDT-based policies on the quality of patient care in relatively optimal conditions. An analysis of existing dataset from a pragmatic trial from Uganda is used to evaluate the effect of an RDT-based policy on the quality of patient care and on incremental cost in routine practice in Uganda. A Health Facility Assessment (HFA) is used to determine the quality of current practice and adequacy of support services when RDT-based guidelines are rolled out to health services in a district as a whole as routine practice.

6.2 Summary of results

Major findings from the 3 studies are summarised in Table 49, which also specifies the indicators assessed in 2 or more studies, and those that were study-specific.

Table 49: Summary of main findings from chapters 3, 4 and 5

Variables	**Indicators**	**Systematic review** (RCTs)	**Pragmatic trial**	**Survey of current practice**
Quality of practice	*Adequacy of clinical assessment*	No data	No data	Inadequate
	Use of RDTs	No data	Most cases (97%) were tested with RDTs prior to any treatment.	"Low" in low transmission district; "high" in high transmission district Quality of use "low" in both high and low transmission districts.*
	RDT-positive cases getting antimalarials	98% to 100% received antimalarials (3 studies)	98% to 100% received antimalarials (all transmission settings)	All cases with positive RDTs received antimalarials in both high and low transmission districts
	RDT-negative cases getting antimalarials	0.4% received antimalarials in 1 study (high adherence) 40% to 81% received antimalarials in 3 studies (poor adherence)	0.5% received antimalarials in all transmission settings (high adherence)	None of the cases with negative RDT results received antimalarials in both high and low transmission districts
	Advice on medications	No data	No data	Adequate in low prevalence district; inadequate in high prevalence district
Outcomes	*Patients receiving antimalarials*	Wide variability in results: number declined significantly where adherence to RDT results was high. No difference in prescribing where adherence to RDT results was low.	Significant decline in proportion of patients prescribed antimalarials in all transmission areas. Highest in the low prevalence district and lowest in the high transmission district— effect higher among older (>5 years old) patients	Not measured
	Patients receiving antibiotics	2 studies: increased in one study. No difference detected in another	Declined by 5% in high transmission setting. No significant change in medium and low transmission settings	Not assessed

**HWs failed to follow some critical steps in carrying out the test*

Variables	Indicators	**Systematic review** (RCTs)	**Effectiveness trial**	**Survey of current practice**
Outcomes (continued)	*Patients unwell 4 – 7 days after treatment*	No difference detected between RDT and clinical diagnosis arms	No difference detected between RDT and clinical diagnosis arms	Not assessed
	Cost	No data	Significant decline in low and medium transmission settings only. Significant declines in older (> 5yreas old) patients only	Not assessed
	Malaria cases missed	No difference (1 study)	Corresponds to RDT false negative error rates. High in low prevalence areas.	Not assessed
	Non-malarial cases prescribed antimalarials	Significant declines, but remains high at intervention sites (1 study)	Corresponds with RDT false positive error rates. High in high transmission setting; low in low transmission setting.	Not assessed
System readiness	*Essential inputs*	No data	No data	RDTs: inadequate supply—out of stock for 8 months in both districts. Essential drugs: inadequate supply in both districts Guidelines: inadequate presence in the health facilities Equipment: inadequate in both districts Staff: inadequate in both districts
	Support services	No data	No data	In-service training: inadequate in low prevalence district; adequate in high prevalence district. No refresher training in 2-3 years. Support supervision was adequate in both settings

6.2.1 Cochrane review

Do carefully controlled trials show that treatment policies for treating fever which are based on RDT results lead to better quality of care than treatment based on clinical judgement?

Four trials met the inclusion criteria (n=9545): 2 individually randomised and 2 cluster.

a) Quality of practice

Two indicators were used to assess the quality of practice in controlled conditions; namely, (a) proportion of RDT-positive patients for whom antimalarials were prescribed (b) the proportion of RDT-negative patients for whom antimalarials were prescribed. HWs were expected to perform each of the two tasks correctly at least 95% of the time. Otherwise the quality of practice or HW performance was judged to be inadequate [104, 106, 127, 128, 140]. Only 3 studies provided data for the first indicator [34-36]. HW response was adequate in all three trials: 98% to 100% of cases with positive RDT results were prescribed antimalarials. HW response varied widely with regard to the second indicator. HW performance was adequate in only one trial in which antimalarials were prescribed to 0.4% of patients with negative RDT results [37]. HW performance was inadequate in 3 trials: patients with negative RDTs for whom antimalarials were prescribed varied from 40% to as many as 81% [34-36]. Therefore, the quality of practice was adequate in one trial only, in which HWs withheld antimalarials from 99.6% of patients who had negative RDTs.

b) Primary outcomes

Fewer patients were prescribed antimalarials in the intervention (RDT) facilities than in the control (clinical diagnosis) facilities (combined RR = 0.71, 95% CI 0.56 to 0.90, 4 trials, n = 9327).

- Estimates of effects of the intervention varied considerably across the 4 studies (Chi^2 = 175.78, df = 3, (P < 0.00001), I^2 = 98%).

- Studies detected substantial decline in prescribing of antimalarials at intervention HCs where HW adherence was high (RR = 0.23, 95% CI 0.14 to 0.38) [37].

- Trials with low HW adherence at intervention HCs detected no significant difference in prescribing of antimalarials between the two arms (combined RR = 0.85, 95% CI: 0.66 to 1.08) [34-36].

- Variability in age group did not influence the effect of the intervention (test for sub group difference: $Chi^2 = 0.44$, df = 1 (P = 0.51), $I^2 = 0\%$).

Studies detected no significant difference between RDT and clinical diagnosis arms for patients who were still unwell between days 4 - 7 (RR = 0.89, 95% CI: 0.68 to 1.28).

c) Secondary outcomes

Targeting of antimalarials

- One study detected a significant reduction in the intervention arm in non-malaria fevers (reference microscopy negative patients) prescribed anti-malarials (RR = 0.60, 95% CI: 0.57 to 0.64).
- The same study detected no difference between RDT and clinical diagnosis arms in true malaria cases (reference slide positive patients) not receiving anti-malarials (RR = 1.21, 95% CI: 0.64 to 2.23)[36].

Antibiotics prescribing

No significant difference was detected between the 2 arms in the amount of antibiotics which was prescribed (RR = 1.03, 95% CI: 0.97 to 1.10; 2 studies) [34, 36].

6.2.2 Effectiveness trial

Does the evidence from a pragmatic trial in Uganda show improved quality of patient care when treatment policy for fever based on RDT results is applied to a large population through routine health services?

a) Quality of practice

This trial took place in low, medium and high transmission settings. Judgement of HW performance was based on assessment of 3 tasks: (a) use of RDTs, (b) prescribing of antimalarials in response to positive RDT results, and (c) prescribing of antimalarials in response to negative RDT results. HWs were expected to perform each of the 3 tasks correctly at least 95% of the time. Otherwise the quality of practice or HW performance was judged to be inadequate [104, 106, 127, 128, 140].

RDTs were used in at least 97% of participants across all intervention sites. At least 98% of patients with positive RDT results received antimalarials. Less than 0.7% of participants with negative RDT results were prescribed any antimalarials. Therefore, HW performance was adequate in all malarial transmission settings, and in all age groups.

b) Primary outcomes

After introducing RDT-supported guidelines,

i) **prescribing of antimalarials** declined significantly in all transmission settings, as follows:

- high transmission by 22.1% (95% CI: -24.6% to -19.6%)
- medium transmission by 48.9% (95% CI: -52.5% to -45.5%)
- low transmission by 60.2% (95% CI: -64.1% to -56.3%).

ii) **Prescribing of antibiotics**

The marginal change in prescribing of antibiotics was as follows

- high transmission: -5.0% (95% CI: -8.7% to -1.3%)
- medium transmission: -2.7% (95% CI: -7.0% to 1.5%)
- low transmission: -4.6% (95% CI: -9.7% to 0.5%)

iii) Healthcare cost

- high transmission: no significant change detected: US$ +0.02 (95% CI: US$ -0.97 to US$+1.06).
- medium transmission: reduced by 17.7%, amounting to a cost saving of US$ 0.33 per case of fever (95% CI: US$ -0.54 to US$-0.12). Reduced in older children and adults (5+ years) only.
- low transmission: reduced by 24.5%, amounting to a cost-saving of US$ 0.50 per case of fever (95% CI: US$ -0.69 to US$-0.31). Reduced in older children and adults (5+ years only).

iv) Clinical outcomes

No difference was detected between RDT and clinical diagnosis arms in the number of patients still unwell at day 5

- high transmission: RR = 1.02 (0.97 to 1.06)
- medium transmission: RR = 1.00 (0.97 to 1.04)
- low transmission: RR = 1.02 (0.98 to 1.07)

c) Secondary outcomes (targeting of antimalarials)

After introducing RDTs at intervention HCs some non-malarial (reference slide-negative) patients received anti-malarials at intervention HCs despite high HW adherence to guideline. The number in high, medium and low transmission settings was 43.1%, 15% and 11% respectively. These numbers correspond with the RDT α-error rates at the respective sites, namely 43.9%, 15.2% and 10.5% respectively.

On the other hand, some malarial (slide positive) fever patients were missed at the intervention sites after introducing RDTs. The proportion missed at high, medium and low transmission sites was 1.9%, 5.9% and 17.2% respectively. These figures closely correspond with the RDT false-negative error rates at the respective sites, namely 2.0%, 4.5% and 17.2% respectively.

6.2.3 Survey of current practice

Are all components of an RDT-based guideline being implemented to acceptable standards in the districts in Uganda where RDT-based policy has been rolled out for routine implementation? How adequate are essential health systems inputs and support services?

a) Quality of practice

A survey of current practice took place one year after rolling out the RDT-supported policy for fever for small-scale implementation. The assessment took place in 2 districts representing low and high malaria transmission intensity respectively. All components of care specified in the RDT-based guideline for Uganda were assessed, consisting of a total of 25 tasks. The study used the LQAS technique to assess the quality of practice in a district as a whole, based on the quality of guideline implementation at constituent HCs. A district was judged as high performance if at least 80% of the sampled HCs in it implemented the RDT-based guideline adequately. An HC was judged as high performance if the most experienced HW at the HC performed each of the 25 tasks correctly at least 95% of the time. These standards have been used in several LQAS-based surveys [104, 106, 127, 128] and were discussed with officials of the Ministry of Health and the districts prior to application.

None of the HWs met the performance threshold for all the tasks observed. Therefore, the quality of practice was judged to be inadequate in both high and low transmission districts. Clinical assessment was particularly incomplete in both transmission areas. More surprisingly, RDT was requested in only 29.2% of consultations in an area with low malaria prevalence, compared with 91.7% of consultations in an area with high malaria prevalence. In both high and low transmission districts, HWs committed errors in several steps involved in performing an RDT test. Most critically, RDT results were read hurriedly, often within 2 minutes of the recommended 15 minutes waiting period. The quality of prescribing of antimalarials was low in low prevalence area, mainly due to low utilisation of RDTs. In high prevalence area, prescribing of antimalarials matched RDT results most of the time. Therefore, prescribing of antimalarials was judged as being appropriate in high transmission setting despite the fact that the quality of clinical assessment and RDT test was poor.

b) Essential inputs and support services

None of the districts had adequate stock of all essential drugs required for effective implementation of the RDT-supported guideline—on the survey day or during the

preceding 6 months. RDTs were out-of-stock in both districts, both on the day of the visit and throughout the 6 months preceding the survey. ACTs were available in adequate amount on survey day in both districts. In the preceding 6 months, supply of ACTs was erratic in an area with high malaria prevalence, but adequate in an area with low malaria prevalence. The most problematic drug items were cotrimoxazole (for treating pneumonia or ARI) and CIRPO (dysentery)—both on the survey day and during the 6 months prior. Relevant guidelines and essential equipment were also in insufficient number in both districts. Although none of the districts had adequate number of all the recommended staff, both districts had adequate number of professional staff. In-service training on use of RDT-based guidelines was adequate in high prevalence district but low in low prevalence district. Support supervision was adequate in both districts.

6.3 General discussion

The basis for adopting a new technology should be the evidence that it is more effective than the intervention currently or previously in use, and that, once scaled-up, it can be implemented sufficiently [40, 41]. Effectiveness is achieved only if an efficacious intervention is implemented sufficiently and the target audience is receptive to or adheres to it. Accordingly, the main questions regarding utility of RDT-based guidelines has been whether they can be implemented sufficiently in routine practice and whether their use in routine practice can lead to better outcomes than treatment based on clinical judgement [14, 26, 34]. There are indeed concerns that RDT-based treatment may lead to increased morbidity due to missed diagnosis of malaria cases [26, 34].

6.3.1 Is the intervention efficacious and effective?

a) Effect on clinical outcomes

The review in chapter 3 included two trials which evaluated effects of RDT-based guidelines on clinical outcomes under experimental conditions. None of the 2 studies detected a significant difference in clinical outcomes between RDT and clinical diagnosis arms. Additionally, the analysis of data from effectiveness trial in chapter 4 also detected no statistically significant difference in clinical outcomes between RDT-based treatment group and clinical diagnosis group—even in a low transmission area where more than 17% of malarial fevers were missed by RDTs. According to the literature review presented in section 2.5, failure to demonstrate differences in outcomes between comparison groups might be due to (a) inadequate implementation of an intervention, (b) inefficacy or weak causal relationship between the intervention and clinical outcomes, (c) low level of acceptance by the target population, and (d) methodological limitations. A combination of these factors may have precluded demonstration of effects of RDT-based guidelines both under experimental and semi-experimental conditions.

Were interventions poorly implemented?

We do not know if all components of RDT-guidelines were implemented sufficiently in the trials analysed in chapters 3 and 4. On account of the indicators assessed—i.e. HWs' use of RDTs and prescribing of antimalarials relative to RDT results—we can construe that RDT-based guidelines were poorly implemented in the trial by Bisoffi and colleagues in Burkina Faso (chapter 3) [34]. The authors acknowledge that inadequate

HW adherence to RDT results may have undermined the ability of this trial to detect any significant difference in clinical outcomes between the comparison groups [34]. On account of the same measures, we can conclude that RDT-based guidelines were sufficiently implemented in Zambian children in the trial by Yeboah-Antwi and colleagues (chapter 3) [37]; and in the effectiveness trial from Uganda (chapter 4). However, these latter 2 trials also detected no significant difference in clinical outcomes between the comparison groups. Therefore, if the indicators assessed were indicative of the overall level of guideline implementation, then insufficient implementation would not fully explain failure of the trials to detect differences in clinical outcomes. In reality, we do not know if guideline implementation was adequate in both efficacy and effectiveness trials, since measurement was based on a limited set of indicators, which are known to over-estimate overall performance of HWs. It is possible that other vital components of the guideline have been poorly implemented, thereby hampering demonstration of effects in both studies.

Is it weak causal linkage?

Efficacy trials are concerned with demonstrating whether or not an intervention works, and how and why it works [40, 97]. The causal relationship between a clinical guideline and clinical outcomes is complex [73, 74]. A clinical guideline targets change in HW performance. Therefore, its most direct effect consists of the proficiency with which HWs execute clinical tasks. Implementation of a clinical guideline is linked to clinical outcomes indirectly through its effect on the quality of the process of care. Additionally, literature suggests that change in clinical outcomes is determined by several other factors which may be independent of the change in HW performance, such as the efficacy of medications received, patients' adherence to medications, and antecedent factors such patients' demographics [83, 86]. Therefore, adequate implementation of a guideline may not translate into significant differences in clinical outcomes because of the effects of intermediary factors. Conversely, improvement in clinical outcomes does not necessarily indicate that a guideline is efficacious or effective.

Furthermore, literature indicates that in a multifaceted intervention like an RDT-based guideline, different components may be linked to different outcome attributes [73, 74]. For example, we know that RDT-directed prescribing may save antimalarials cost. At the same time, it may result in increased morbidity due to increased risk of some malarial cases missing antimalarials [26]. For this reason, it is presumed that

improvement in clinical outcomes may result mainly through identification of parasite negative patients in whom alternative diagnosis can be sought [19, 70-72], through adequate clinical assessment [138, 155, 157, 166].

Like several observational and quasi-experimental studies evaluating the utility of RDT-based guidelines [32, 120, 121, 151, 152], the trials analysed in chapters 3 and 4 did not evaluate implementation of several vital components of the guideline—notably clinical assessment and proficiency in the use of RDTs, which can significantly influence the quality of diagnoses. Therefore, we do not know for certain, the components of the guideline, which if adequately implemented, may have led to significant improvement in clinical outcomes. In essence, the clinical outcome attributes assessed may not be valid measures of efficacy and effectiveness of RDT-based guidelines.

Is it poor adherence by patients?

Patients who are denied antimalarials on account of their negative RDT results may obtain antimalarials from other sources if they believe that malaria is the cause of their illnesses. Further patients who do not improve on the current treatment may also seek further treatment elsewhere and may not be accounted for in evaluation of outcomes. These behaviours can mask differences in clinical outcomes between the comparison interventions; and might explain the failure to detect a difference in clinical outcomes in Mubende (low transmission) where more than 17% of malarial patients were missed by the use of RDTs.

Is it due to methodological limitations?

Conditions for demonstrating causal relationship did not pertain fully in the trials which evaluated clinical outcomes in chapter 3. For example, allocations were not concealed in all the trials and this could have led to selection bias [34, 37]. Further, in both trials, study participants, providers and outcomes assessors were not blinded to the intervention, test results, diagnoses, treatment prescribed and outcomes assessed [34, 37]. Lack of allocation concealment and/or blinding is known to bias results towards positive outcomes rather than null effects [99]. Therefore, the results of the trials in chapter 3 and the analysis in chapter 4 may not have been due to these methodological limitations.

Further, the criteria used in selecting studies for inclusion in the review in chapter 3 may have been biased towards trials with no significant effects. On the other hand, they may

have been biased against trials considered to have sub-optimal designs, but which may have had large effect size [73].

Therefore, we are unclear whether or not use of RDT-based guidelines has any effects on clinical outcomes; or the components of the guidelines with the most significant effects on clinical outcomes. More carefully controlled trials which also assess implementation of all components of RDT-based guidelines may help identify any causal linkage between various components of the RDT-based guideline and various outcome indicators.

b) Effect on prescribing and cost

The analysis in chapter 4 shows that use of RDT-supported guidelines in the treatment of fevers can lead to significant savings in drug costs in older children and adults (5+ years), and in low to medium transmission settings. Costs consisted of the values of resources used (RDTs, antimalarials, antibiotics and analgesics) during an effectiveness trial which was characterised by high utilisation of RDTs and near perfect HW adherence to RDT results. The analysis detected no significant change in the values of antibiotics and analgesics used before and after introducing RDTs. Effectively, cost-savings were accrued by averting inappropriate use of antimalarials.

These findings are consistent with those from several economic evaluations which are based on assumption of perfect HW adherence with regards to use of RDTs and response to RDT results [12, 13, 72, 140, 144]. However, our evidence is inconsistent with an effectiveness trial from Kenya where inadequate adherence to RDT results in a low transmission setting was found to be associated with a substantial increase in the cost of diagnosing and treating a case of fever.

The analyses in chapter 3 and several observational studies indicate that HWs adherence to RDT results and prescribing of antimalarials can vary even under carefully controlled conditions [32-37]. Additionally, evidence from chapters 3 and 4 of this thesis also shows that use of RDT-based guideline may have no effect on prescribing of antibiotics [34], or it may increase its usage [36]. Further, two effectiveness trials from Tanzania also indicate that RDT-directed treatment may substantially increase prescribing of antibiotics [121, 156]. Increase in prescribing of antibiotics wound discount cost-savings accrued by averting irrational prescribing of antimalarials. Therefore, in light of the unpredictable HWs response to RDT results and unclear effect of the intervention on

prescribing of antibiotics, we can not predict with certainty how application of an RDT-based guideline might impact on the magnitude of drug cost saved in routine practice.

These results need to be interpreted with caution. We know that data on prescribing are correlated, and potential design effects may have reduced the power of the studies analysed to detect significant and/or larger effects of the intervention on prescribing and costs, and may have resulted in over-precise estimates [99, 102]. Further, the analysis of costs captures the values of biomedical consumables only, and therefore underestimates costs. However, studies which have incorporated programmatic costs in their analyses (e.g. staff time, training cost, and supervision cost) have also reported significant cost-savings in low and medium transmission settings [72, 140]. It is likely that this analysis would have nevertheless detected significant cost-savings in older children and adults, and low to medium transmission areas, were it to include programmatic costs.

6.3.2 Can it be implemented sufficiently in routine practice?

The results of the survey of current practice in Uganda suggests that once scaled-up for use in routine practice with limited supervision, none of the components of an RDT-based guideline might be implemented to acceptable standards. Implementation of the diagnostic components, namely assessment and steps in performing an RDT, could be particularly problematic.

To my knowledge, this is the first survey of an RDT-based guideline which attempts to evaluate the implementation of all components of the guideline, alongside an assessment of essential health systems factors. The findings are consistent with those from previous evaluations of IMCI-oriented guidelines in Africa, which reveal that the quality of guideline implementation is often much poorer than is portrayed by indicators of quality of prescribing alone [33, 90-93, 109, 110, 112, 113, 115, 117, 146, 150]. Performance scores for clinical assessment in the context of IMCI has been found to vary widely across studies [90, 92, 93, 113], from a low of 27.3% in Benin [115] to a high of 86.9% in Morocco [90]. The pooled performance scores for clinical assessment in the 2 districts surveyed—16% in Kisoro and 26% in Gulu—appear to be on the poorer side compared to the performance scores reported in several surveys of IMCI guidelines (see pooled scores in Appendix 20). It is acknowledged that meaningful comparison of our findings with those of previous surveys is hampered by variability in contexts, the small

sample size in our survey, number of tasks assessed in various surveys, and design effects.

Adequate clinical assessment is considered vital in improving the quality of diagnoses, and probably clinical outcomes, especially in fever cases who are parasite negative [138]. For example, a sub-analysis of diagnostic pathways in an evaluation study of an IMCI guideline in Benin revealed that some of the diagnoses upon which treatment was based were false because of errors related to inadequate clinical assessment of patients; yet treatment matched diagnoses all the time in those circumstances [149]. The sample size involved in the sub-analysis was small. However, an evaluation of IMCI guideline in Ghana which involved follow-up of patients and evaluation of outcomes found no association between the extent of clinical assessment and clinical outcomes [118]. Additionally, an assessment of an IMCI guideline in Benin in which patients were re-evaluated by expert research staff found out that HWs often prescribed effective treatment despite deviation from guidelines [115]. These finding raise questions about the effectiveness of the current clinical criteria in guiding treatment in a child with fever. The accuracy of symptoms-based diagnoses and treatment in HCs manned mainly by nurses and paraprofessional staff is bound to vary. Inadequate clinical assessment may exacerbate the ability of the staff to make accurate working diagnoses.

We know from economic models (chapter 2) and the analyses in chapter 4 that use of RDTs can significantly save cost of medical consumables in low to medium transmission areas and in older children and adults only. Contrary to expectations, use of RDTs was low in Kisoro (low transmission district) and high in Gulu (high transmission district). We do not know if this behaviour pattern is reflective of low transmission and high transmission regions in Uganda generally. Additionally, we do not know if HWs in both districts would have behaved differently towards older patients.

A similar pattern in the use of RDTs in fever case management has been reported in Kenya, where it was found to be associated with a substantial increase in diagnosis and treatment cost in a low transmission setting [33]. If low utilisation of RDTs as observed in Kisoro district in Uganda, and in Kenya is indicative of HWs' behaviour in low transmission areas as a whole, then RDT-based treatment policies for fever may lose their economic advantage over treatment based on clinical diagnosis, even in low and moderate transmission areas.

More disturbing was the finding in both districts that, when RDTs were used, HWs committed errors in several steps involved in performing the test. Most critically, HWs read the results hurriedly, often within 2 minutes of the recommended 15 minutes waiting period. This could be a signal that the quality of RDTs in routine practice where there is limited supervision could be poorer than is known from findings from trials.

Already we know from literature that, even under experimental conditions, RDTs often fail to achieve the desired sensitivity of 95% at parasite density of ≥ 100 parasite/100 μL of blood [20, 22, 65, 135]. Errors in using RDTs could significantly exacerbate the sensitivity of RDTs, increasing the risk of morbidity and mortality from missed malaria cases, especially in low transmission areas and in children [26, 138]. Although evidence from efficacy studies (chapter 3) suggests that the risk of missed malaria cases is small and similar in both RDT and clinical diagnosis arms, evidence from the effectiveness trial (chapter 4) indicates that this risk can be substantial. More than 17% of slide positive patients were not detected by RDTs at the site in low transmission setting, which might indicate poor proficiency in carrying out an RDT test in a trial setting.

Several health systems factors are known to positively influence HWs adherence to guidelines to various extents (section 2.10). Support services such as in-service training and supervision are generally associated with improved HW adherence to guidelines [90, 93, 108, 112, 117, 146]. Availability of guidelines has been found to have mixed effects on guideline implementation [93, 112, 114, 146, 150], while availability of essential drugs appears to have no significant influence on guideline implementation [93, 112, 150].

The health systems in the two districts assessed were characterised by prolonged shortage of RDTs, shortage of job aids and lack of refresher trainings, all of which may explain the poor performance quality in both districts. Human capital theory (sub-section 2.3.2) underlines the importance of consistency of use of a new technology in order to maintain proficiency over time [28, 31, 41, 42, 46], especially if its use requires the acquisition of new skills [41]. An LQAS assessment of family planning services in Kenya found that HWs who continued to use their skills had better skills-retention level than those who were no longer offering the same services at the time of the survey [106]. RDTs are a central component of the new intervention. Prolonged disruption in the use of RDTs could have resulted in loss of interest and loss of skills in its use in

both districts [60, 61, 106, 183, 184]. Inadequate exposure to in-service training may have exacerbated the performance quality in Kisoro district (low transmission area).

In essence, introduction and sufficient implementation of the intervention requires strengthening of the health system by ensuring regular supplies of RDTs and job aids, and by providing sufficient in-service training and support supervision. Additionally, because of the likelihood of inter-lot variability, the WHO strongly recommends that countries should develop mechanisms for lot-testing at national level, and for regular random testing the level of use [65]

6.3.3 Has the intervention been evaluated sufficiently?

The technology under evaluation is a complex policy, of which use of RDTs is just one component. The current debate, and the focus of this thesis, is whether there is sufficient evidence to show that the policy can be effective in clinical settings. Extensive laboratory-based [65] and field-based trials [20, 22, 135] have been carried out, and are ongoing, aimed at improving the quality of manufacturing and performance of assays. However, limited evaluation has been undertaken to determine utility of *RDT-based guidelines* in the field.

As shown in this thesis, both observational [32, 33, 120] and randomised trials [34-37] examining utility of RDT-based guidelines have focused mainly on the effect of the intervention on prescribing of antimalarials. Although variable, evidence from these studies indicate that RDT-based guidelines can be effective in reducing use of antimalarials and treatment cost in older children and adults in low and medium transmission areas—as long as HWs use RDTs and prescribe antimalarials according to the results. Qualitative research could provide an in-depth knowledge into context-specific factors responsible for variability in HW use of RDTs and adherence to RDT results. Fewer trials [34, 36] and observational studies [121, 156] have evaluated effects of the intervention on prescribing of antibiotics. The effectiveness trial analysed in chapter 4 is the most recent. They show that the intervention may not have any effect on prescribing of antibiotics or they may increase it.

Studies examining clinical outcomes have been more limited: 2 trials analysed in chapter 3 [34, 37], the effectiveness trial analysed in chapter 4 and one observational study [121]. Only the observational study detected a marginal but statistically significant difference in

persistence of fever (2.5% in RDT arm, and 5% in clinical diagnosis arm) [121]. Therefore, effects of the intervention on clinical outcomes need further examining, because this is currently the most contentious aspect of the policy, and given the possibility of high false positive error rate in routine practice as shown in chapter 4 and as implied in chapter 5. Studies evaluating effects of the intervention on clinical outcomes need to include an assessment of implementation of all components of the guideline. When incorporated into a model, variation in implementation of various components of a guideline may be used to predict specific outcomes attribute [85]. Given the limitations of RCTs in assessing complex public health interventions, quasi-experimental can be useful for attributing effects to a multi-component intervention. Clinical outcomes might not be a valid indicator of improvement in the quality of practice. However, there is need to demonstrate that RDT-based guidelines are not inferior to clinical diagnosis in terms of their effects on clinical outcomes.

If the decision for wide-scale application of an RDT-based guideline is to be based on clinical outcomes, then we can conclude that, presently, there is insufficient evidence to support that decision. It is recognised, however, that RDT-based guidelines have already being rolled out in some countries [25], despite the insufficient evidence of their effects on clinical outcomes. The bases for those decisions are unclear. If policy makers decide to disseminate the policy on the basis of economic potentials, then the evidence from literature and this thesis (chapter 4) suggest that it would be prudent to restrict their application to low and medium transmission settings and in older children.

6.3.4 What this research contributes

Use of theory-based model: This thesis has shown how the diffusion of innovation model can be used to provide a framework for systematically and comprehensively reviewing a complex healthcare technology. By applying the diffusion of innovation model to the evaluation of the RDT-supported policy for treating fever, this thesis has provided different types of information which policy makers might require in making adoption and implementation decisions, including an indication of:

- how the quality of implementation of the guideline and outcomes might vary within and between carefully conditions versus in routine practice
- key health systems barriers that might need to be addressed should the decision be taken to scale-up implementation of the innovation

Comprehensive assessment of guideline implementation: The thesis assessed the quality of delivery of all components of the guideline one year after being rolled out for small scale implementation. Therefore, it provides a comprehensive picture of what the quality of practice might be after scaling-up the implementation in routine practice. It also shows the quality of all the specific components of the healthcare process, including the diagnostic procedures.

- It identifies critical components of the guidelines which are being implemented poorly, and which need to be improved through specific interventions prior to wide-scale implementation.
- It indicates that assessments based on the use of RDTs and appropriateness of antimalarials alone may over-estimate the overall quality of implementation of the guideline [109, 113].
- To my knowledge, this is the first study in Africa that provides a comprehensive picture of the quality of implementation of RDT-based guidelines, alongside a systematic assessment of the state of essential health systems inputs and support services which are required for routine implementation.

Use of LQAS method to prioritise areas for improvement: In this thesis I used Lot Quality Assurance Sampling (LQAS) method to assess whether districts are adequately embracing the RDT guidelines in routine clinical care. LQAS optimises the identification of extreme performance: worst of worst, which require priority attention

and/or best of best, so that resources are not wasted. Through the application of LQAS techniques, this thesis has been able to specify components of the guideline which require urgent attention [106].

Effect of multiple factors on effects of intervention: We know that several factors interact to influence the effect of the intervention (section 2.8.3). The existing studies either do not describe the prevalence of malaria in the setting studied, or they present results in a combined form [34-37]. In this thesis, I used patient-level data to show how variability in malaria prevalence, age group of the target population and HW adherence all combine to influence the effects of the intervention. This information can be useful in deciding whether and how to target the policy to different epidemiological settings and population groups in order to optimise its benefits.

6.4 General Conclusions

6.4.1 Conclusions about the research questions

The main questions this thesis set out to answer were whether RDT-based guidelines can be implemented to acceptable standards in routine practice and whether their use in routine practice can lead to better quality of care in patients with fever—in terms of prescribing, cost and clinical outcomes—relative to symptoms-based treatment. Effectiveness can be achieved only if an efficacious intervention is implemented sufficiently and that the target audience adheres to it. The thesis provides evidence on efficacy of the intervention in order to aid interpretation of effectiveness results.

- Evidence from the review of RCTs, the effectiveness trail from Uganda and other studies show that antimalarial use is lower in low, medium and high malaria transmission settings when RDTs are used to guide treatment of fever instead of presumptive treatment. This research shows that this results in savings from drug costs in older children and adults with fever in low and medium transmission areas.

- Health worker adherence and use of RDTs results vary widely, even under experimental conditions. Therefore, the amount of antimalarials and costs saved by RDT-based guidelines instead of presumptive treatment of fever may vary widely in routine practice with limited supervision.

- The review, the effectiveness trial from Uganda and other studies show that RDT-based policies may have no effect on prescribing of antibiotics or it may increase it. Therefore, in some settings, the amount of cost-savings accrued by averting irrational prescribing of antimalarials may be discounted by an increase in prescribing of antibiotics.

- This study combined with others, still does not confirm whether or not use of RDT-based guidelines has any effects on clinical outcomes. We do not know if this is due to weak causal linkage between the intervention and outcomes, insufficient implementation of vital components of the guideline or due to methodological limitations. Further, the thesis has shown that RDT-guided treatment may be associated with increased risk of missed diagnosis of malaria in routine practice.

However, we do not know if this can result in significant increase in the risk of morbidity or mortality in routine practice.

- A case study of Uganda suggests that when delivered through routine services, none of the components of an RDT-based guideline is implemented to acceptable standards. The diagnostic components, namely assessment and steps in performing an RDT, are particularly poorly implemented. Therefore, while the policy can save drug costs through the use of RDTs, it may not improve the quality of diagnoses, especially among parasite negative fevers.

- The Uganda case study also shows the policy is being rolled out in a health system which lacks essential inputs and support services required for its effective implementation. This may exacerbate the quality of guideline implementation in routine practice

Therefore, if restricted to older children and adults (5+ years) in low and medium transmission settings, RDT-based policies for fever can significantly reduce usage of antimalarials and save drug cost in routine practice where presumptive treatment is the norm. However, because of the unpredictable HWs response to RDT results and variable effect of the intervention on prescribing of antibiotics, we can not predict with certainty how application of an RDT-based guideline might impact on the magnitude of cost savings in routine practice. There is insufficient evidence to suggest that the policy is superior to presumptive treatment of fever in terms of effects on clinical outcomes. HW proficiency in carrying out clinical assessment, RDTs and counselling is poor when the policy is delivered through routine health services characterised by insufficient supplies and training.

6.4.2 Implications for practice and policy

The evidence in this thesis and other research is insufficient to support a wide-scale implementation of RDT-based policy for fever on account of its effects on clinical outcomes. If policy makers decide to scale-up application of the policy on account of its economic potentials, then the evidence from literature and this thesis (chapter 4) suggests that it is worthwhile to restrict their application to low and medium transmission settings and in older children and adults.

Varying diagnostic strategies for malaria by transmission settings can pose practical challenges in implementation [26, 70]. Therefore, from a practical standpoint, the policy could be implemented across the whole country as long as its use is limited to 5+ years old. Cost-effectiveness models which incorporate longer-term and broader outcome measures such as the DALY, seem to favour use of RDT-based guidelines even in high transmission settings if their use is restricted to 5+ years old (See 2.8.7) [72, 140, 144].

Improvements are required in implementation of all components of the guideline. Particular emphases are required on clinical assessment, criteria for selecting patients for RDTs, steps in performing an RDT and response to RDT test.

- Without improvements in clinical assessment, the quality of diagnosis in parasite-negative individuals cannot be improved. Emphasis on the use of RDTs and appropriate prescribing of antimalarials alone may create the impression that the new policy is only a tool for preserving antimalarials.
- Without improvements in the steps involved in carrying out an RDT in routine practice, several parasite-positive individuals might be missed, resulting in increased morbidity, especially in low prevalence areas where use of the guideline is recommended.
- If HWs do not use RDTs as required, or if they do not prescribe antimalarials according to the results, then the economic advantage of RDT-based guidelines over presumptive treatment is lost even in low to moderate transmission areas.
- These improvements can be achieved through additional interventions, such as refresher trainings, regular support supervision, and provision of job aids and regular supply of RDT kits.

HWs need to perform a minimum volume of tasks with the aid of an innovation in order to maintain an acceptable performance quality. Therefore, it is imperative for health authorities to main an adequate quantity of RDT kits and guidelines at health facilities at all times so that HWs can implement RDT-based guidelines consistently.

6.4.3 Implications for research

a) There is need for in-depth, context-specific, qualitative research to investigate why HW vary in their use of RDT-based guidelines. Qualitative methods can reveal context-specific factors underlying variability in HW use of RDT-based

guidelines. This information can be useful in tailoring interventions aimed at inducing or maintaining behaviour change in users of the technology.

b) Researchers evaluating implementation of RDT-based guidelines should use multiple indicators that enable an assessment of all components of the guidelines. Methods which assess all components of the guideline can identify elements which are poorly implemented. Use of LQAS methods can help identify components that require urgent attention.

c) The evidence from this thesis is unclear about the effects of RDT-based policies on clinical outcomes. More outcomes data are required from randomized trials in order to inform a more robust conclusion about the effects of the intervention on clinical outcomes. Quasi-experimental designs which allow the implementation of various components of the guideline to be assessed may permit attribution of effects to specific components of the guideline.

Appendices

Appendix 1 Formula for binomial distribution

$$P_a = \frac{n!}{a!\,(n-a)!}\, p^a x q^{n-a}$$

where:

P_a	=	the probability of selecting "a" items of desired quality in a sample of "n" items
p	=	the desired performance standard (e.g. 80% of items of desired quality)
q	=	the expected proportion of items of undesired quality (1-p)
n	=	sample size
a	=	the number of items of desired quality
n-a	=	number of items of undesired quality (defective items, also referred to as "d")

Adapted from Valadez (1991) [104]

Appendix 2 Recommended staffing for various levels of health facilities in Uganda

Referral level	Staffing
HC I	Village health team
HC II	2 minimum staff: - Enrolled nurse (1) - Enrolled midwife (1) - Nursing assistant (1)
HC III	3 minimum staff: - Senior clinical officer (1) - Clinical officer (1) - Enrolled midwife (2) - Nursing officer-nursing (1) - Enrolled nurse (3) - Laboratory technician (1) - Laboratory assistant (1) - Nursing assistant (3) - Health assistant (1)
HC IV	23 minimum staff - Senior medical officer (1) - Medical officer (1) - Senior nursing officer (2) - Nursing officer-nursing (1) - Nursing officer-midwife (4) - Nursing officer-psychiatry (1) - Enrolled nurse (3) - Enrolled midwife (8) - Nursing assistant (5) - Clinical officer (1) - Clinical officer-eye (1) - Health Inspector (4) - Laboratory technician (2) - Assistant VCO officer (1) - Theatre assistant (2) - Laboratory assistant (2) - Health assistant HA (2)
General hospital	The following are minimum requirements - Medical officers (4) - Public dental officers (2) - Dispensers (2) - Senior nursing officers (5) - Nursing officers-nursing (17) - Nursing officers-midwifery (3) - Enrolled psychiatrics (2) - Enrolled nurse (46) - Enrolled midwife (25) - Nursing assistants (15) - Clinical officers (5) - Radiographers (2) - Orthopaedic officers (2) - Anaesthetic officers (2) - Theatre attendants (2) - Laboratory technicians (2)
Regional referral hospital	No data
National referral hospital	No data

Appendix 3 Detailed Search Strategy for Electronic Databases-Chapter 2

a) Concepts and methods

Search set	Search terms
1	Malaria
2	Transmission
3	Intensity
4	1-3/and
5	Endemicity
6	1 and 5
7	Holoendemic*
8	Hyperendemic*
9	Mesoendemic*
10	Hypoendemic*
11	7-10/or
12	Spleen
13	Parasite
14	12 or 13
15	Rate*
16	14 and 15
17	Effective*
18	Pragmatic
19	Efficac*
20	17 or 18 or 19
21	Trial*
22	20 and 21
23	Demonstration
24	Implementation
25	Process
26	23 or 24 or 25
27	Evaluation
28	26 and 27
29	Quality
30	Care
31	29 and 30
32	Framework
33	Model
34	32 or 33
35	29 and 30 and 34
36	Cluster
37	Design effect
38	Inter-cluster correlation coefficient
39	36 or 37 or 38

b) Malaria rapid diagnostic tests: technology, diagnostic accuracy and performance

Search set	Search terms
1	Malaria
2	Fever
3	Febrile illness
4	1 or 2 or 3
5	Parasitological
6	Parasite-based
7	Definitive
8	Microscop*
9	Rapid diagnostic test*
10	RDT*
11	Assay*
12	5-11/or
13	Diagnos*
14	13
15	Accuracy
16	Sensitivity
17	Specificity
18	Performance
19	15-18/or
20	Requirement*
21	20
22	4 and 12 and 14
23	12 and 19
24	12 and 20

c) Malaria rapid diagnostic tests; economic evaluation

Search set	Search terms
1	Cost*
2	Cost-effective*
3	Cost effective*
4	Effective*
5	Efficacy
6	Efficacious
7	Cost-benefit
8	Cost benefit
9	Utility
10	Cost-utility
11	Cost utility
12	1-11/or
13	Evaluation
14	Analysis
15	Study
16	13 or 14 or 15
17	Malaria
18	Fever
19	Febrile illness
20	17 or 18 or 19
21	Rapid diagnostic test*
22	RDT*
23	21 or 22
24	12 and 16 and 20 and 23

d) Theory of diffusion of innovations (search set 37), implementation of guidelines in Africa (search set 38), effects of intervention on quality of implementation of guidelines in Africa (search set 39)

Search set	Search terms
1	Treatment practice*
2	Management*
3	Prescription behaviour*
4	Adher*
5	Compliance
6	Complying
7	Implement*
8	Adopt*
9	Diffusion
10	1-9/or
11	Clinic*
12	Health*
13	11 or 12
14	Guideline*
15	Protocol*
16	Innovation*
17	New knowledge
18	Evidence
19	14-18/or
20	Theory
21	Theories
22	Model*
23	20 or 21 or 22
24	Intervention*
25	Training
26	Supervision
27	Drug*
28	Medicine*
29	Supplies
30	Equipment
31	Job aid*
32	Support
33	24-32/or
34	Africa
35	Developing countries
36	34 or 35
37	10 and 13 and 19 and 23
38	10 and 13 and 19 and 36
39	10 and 13 and 19 and 33 and 36

e) Use of LQAS methods

Search set	Search terms
1	Lost Quality Assurance Method
2	LQAS
3	1 or 2

Appendix 4 Detailed Search Terms for Electronic Databases-Chapter 3

Search set	Search terms
1	Malaria
2	Fever
3	Febrile illness
4	1 or 2 or 3
5	Rapid diagnostic test*
6	RDT*
7	Presumptive treatment
8	Syndromic approach*
9	Treatment practice*
10	Management*
11	Prescription behaviour*
12	Definite diagnosis
13	5-12/or
14	randomized controlled trial*
15	random allocation
16	double blind method
17	single blind method
18	randomly
19	Clinical trials*
20	14-19/or
21	4 and 13 and 20**
	** Search terms 14-19 will not be applied to CENTRAL

Appendix 5 Characteristics of included studies

a) Ansah 2010 GHA

Methods	Trial design: individually randomised controlled trial Patients evaluated on day 28 after initial contact Reference slides taken on all patients
Participants	Children and adults with suspected malaria Exclusion: pregnancy, illness requiring admission, non-compliance with allocated test/treatment, not living locally Number randomised: 3452 Number analysed for primary outcome (prescribing of antimalarials): 3348 (9.6% loss to follow up)
Interventions	Rapid diagnostic test plus treatment versus clinical diagnosis plus treatment. (A second component examining RDT vs microscopy did not meet our entry criteria) Health workers in both groups received training and held guidelines RDT performed by research team Health workers did not fully comply with guidelines: 49.5% of participants with negative RDT results received antimalarials
Outcomes	Primary Patients treated with anti-malarial treatment who did not have malaria based on reference slides. Secondary 1. Patients not receiving antimalarial treatment who were malaria reference slide positive. 2. Patients prescribed antibiotics 3. Patients with positive reference slides not prescribed anti-malarials 4. Patients correctly treated (i.e. patients that were slide positive and treated with AM + patients that were slide negative and not prescribed anti-malarial treatment
Notes	Country: Ghana RDT: OptiMAL-IT Setting: 3 health centres, of all referral levels Transmission: not indicated Dates: July 2007 to December 2008 Funding: Gates Malaria Partnership

Risk of bias table

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Computer generated blocks of 10
Allocation concealment (selection bias)	Low risk	Numbers placed in sealed opaque envelopes
Incomplete outcome data (attrition bias)	Low risk	Loss to follow-up was low and comparable in both settings (≤3%)
Selective reporting (reporting bias)	Low risk	Reported on all outcomes specified in prospective trial register.
Other bias	Low risk	No other sources of bias identified
Blinding of participants and personnel (performance bias)	Unclear risk	Study participants and staff were aware of allocated tests, the results and prescriptions
Blinding of outcome assessment (detection bias)	Low risk	Blood slides read by 2 independent microscopists blind to study allocation and RDT result.

b) Bisoffi 2009 B'FASO

Methods	Trial design: individually randomised clinical trial, lasting 2 months, 1 month in rainy season, 1 month in dry season
Participants	Number randomised: 2169 (1058 in RDT arm, 1111 in presumptive treatment arm) Number analysed for primary outcomes: (a) prescribing of antimalarials analysis 2169 (0% loss); (b) clinical outcomes: 2095 (3.4% loss) Inclusion: age ≥ 6 years; axillary temperature ≥ 37.5° C Exclusion: severe malaria
Interventions	Intervention: RDT for fever Control: Presumptive treatment Both groups received training and held guidelines RDT performed by research team Health workers did not comply with guidelines most of the time: 81% of participants with negative RDT results received antimalarials
Outcomes	Primary: patients with fever on day 4 Secondary: (1) patients still experiencing other symptoms on day 4; (2) patients given anti-malarials (3) patients given antibiotics
Notes	Country: Burkina Faso RDT: Paracheck Setting: peripheral health centres, Sampling: Convenient selection of health centres to ensure rural/urban representativeness Transmission: Not described Dates: 2006; end of dry season and rainy season Funding: UNIDEA-UNICREDIT Foundation

Risk of bias table

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Computer generated random list
Allocation concealment (selection bias)	Unclear risk	Not indicated
Incomplete outcome data (attrition bias)	Low risk	Loss to follow-up was generally low (95.4% dry season; 97.3% rainy season) but not differentiated by study group Performed available case analysis, although reported to have performed intention-to-treat analysis
Selective reporting (reporting bias)	Low risk	Reported on all study outcomes described in the methodology
Other bias	Low risk	No other sources of bias identified
Blinding of participants and personnel (performance bias)	High risk	Both the study participants and personnel were aware of intervention allocation
Blinding of outcome assessment (detection bias)	High risk	Both the study participants and personnel were aware of the diagnosis made and treatment prescribed

c) Skarbinski 2009 KEN

Methods	Trial design: cluster randomised trial, randomised by health facilities Stratified random selection of facilities, by transmission settings (high/low) and facility type (hospitals, health centres and dispensaries) Took into account a design effect of 2 in sampling Reference slides taken; results not reported Study lasted 4 months
Participants	Inclusion: age ≥ 5 years, irrespective of condition Number of participants randomised: Intervention arm: 799 Number analysed for primary outcome (prescribing of antimalarials): 669
Interventions	Intervention: RDTs for fever patients ≥ 5 years Control: Presumptive treatment of fever Both groups received training and held guidelines RDT performed by health workers Health workers did not fully comply with guidelines: 41% of participants with negative RDT results received antimalarials
Outcomes	Primary outcomes: 1. Fever patients prescribed ACT 2. Microscopy negative patients prescribed ACT Secondary outcomes: 1. RDT negative patients prescribed ACT; and RDT positive patients prescribed ACT 2. Patients prescribed ACT presumptively 3. Patients with known alternative diagnosis receiving ACT
Notes	Country: Kenya RDT: Paracheck Setting: all referral levels of facilities, 60 in total, 30 in each arm Transmission: all transmission levels included Dates: June to September 2006 Funding: USAID

Risk of bias table

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Systematic allocation
Allocation concealment (selection bias)	High risk	Although probabilistic sampling was used in selecting the participating health facilities, the participants had foreknowledge of intervention assignments
Incomplete outcome data (attrition bias)	High risk	Per protocol analysis; loss to follow-up was high (16.3%), more at the intervention facilities (20.2%) than at control facilities (11.2%)
Selective reporting (reporting bias)	Low risk	Reported on all pre-specified outcomes (prescribing of ACT); did not explicitly report on overall anti-malarial prescribing, but the summary data were available for inclusion in the review.
Other bias	Low risk	Baseline imbalance minimised by stratifying facilities by level and randomly selecting within each level Summary data were adjusted for baseline imbalance Loss of whole clusters: not reported Results could be biased towards the null, because some facilities in the comparison arms had RDT
Blinding of participants and personnel	High risk	Both the study participants and personnel were aware of intervention allocation

(performance bias)		
Blinding of outcome assessment (detection bias)	High risk	Both the study participants and personnel were aware of the diagnoses and prescriptions

d) Yeboah-Antwi 2010 ZAM

Methods	Trial design: cluster randomised, by health posts. Pairs matched by distance from health centre, then randomised Patients follow-up and clinical status evaluated 5 – 7 days after initial contact
Participants	Inclusion: Children (6mo – 5 years); presenting with fever with or without other conditions Total enrolled and randomised: 3125 (1017 in the RDT arm and 2108 in the clinical diagnosis arm) Number analysed: (a) analysis 1.1: 3125; (b) analysis 1.2: 3047 (available case analysis)
Interventions	Intervention: RDT-aided algorithm Control: Clinical algorithm Both groups received training and held guidelines RDT performed by health workers; additional interventions provided to increase adherence to guidelines Health workers complied with guidelines most of the time: only 0.4% of participants with negative RDT results received antimalarials
Outcomes	1. Children with fever who received AL 2. Children still experiencing symptoms at follow-up (day 5 – 7)
Notes	Country: Zambia RDT: ICT Malaria Pf (ICT Diagnostics) Setting: Community health posts, manned by community health workers with 6-week training in basic clinical skills, rural and urban Sampling: 42 community health posts Transmission: High prevalence (valley) and low prevalence (plateau) areas Dates: Between December 2007 and November 2008 Funding: Not provided

Risk of bias table

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Authors report random allocation; numbers were generated by random number generator.
Allocation concealment (selection bias)	High risk	Participants were aware before hand, of the diagnostic procedures they were assigned to
Incomplete outcome data (attrition bias)	High risk	Differential loss to follow-up: 4% in the RDT arm vs. 31.4% in the control health posts. Undertook per protocol analysis
Selective reporting (reporting bias)	Low risk	Reported on all pre-specified outcomes
Other bias	Low risk	Recruitment bias was low:-pairs of aid posts were matched by distance then randomised Baseline imbalance: selected clusters were similar and imbalance was adjusted for Loss of whole clusters; no loss was reported
Blinding of participants and personnel (performance bias)	High risk	Both the study participants and personnel were aware of the diagnostic procedures applied
Blinding of outcome assessment (detection bias)	High risk	Both the study participants and personnel were aware of the diagnostic outcomes, and the medications prescribed

Appendix 6 Characteristics of excluded studies

Study	Reason for exclusion
Chinkhumba 2010	Not a randomised trial but a cross-sectional survey, without a comparison group
Faucher 2010	Not randomised trial; and examined the effect of withholding anti-malaria to RDT-positive children rather than comparing RDT-based policy with presumptive treatment
Kyabayinze 2010	Not a randomised trial
Msellem 2009	Not a randomised study (weekly cross-over of intervention)
Reyburn 2007	On account of the comparison-which was policy based on microscopy, rather than presumptive treatment

Appendix 7 Data extraction form

Study ID:
Design: Individual randomised, open label, clinical trial

	RDT-based policy		**Clinical diagnosis**	
Outcomes (and subgroups)	Events	Total	Events	Total
Patients prescribed antimalarials				
≤ 5				
5 +				
All				
Microscopy positive patients not prescribed antimalarials				
≤ 5				
5 +				
All				
Microscopy negative patients prescribed anti malarials				
≤ 5				
5 +				
All				
Patients prescribed antibiotics				
≤ 5				
5 +				
All				
Patients still not well at follow-up (day 4+)				
≤ 5				
5 +				
All				
Comments: Number randomised: Number analysed:				

Appendix 8 Summary statistics for outcomes assessed in Chapter 3

1.0 ALL PATIENTS

		Experiment		Control	
1.1	**Patients prescribed anti-malarials**				
	Study or Subgroup	Events	Total	Events	Total
	Yeboah-Antwi 2010*	27.5%	963	99.1%	2084
	Skarbinski 2009	40.9%	359	54.2%	310
	Ansah 2010	70.0%	1719	92.7%	1723
	Bisoffi 2009	89.3%	1058	87.2%	1111
1.2	**Microscopy positive patients missing anti-malarials**				
	Ansah	3.2%	647	2.7%	633
1.3	**Microscopy negative patients receiving anti-malarials**				
	Ansah	53.9%	1072	90.1%	1090
1.4	**Number prescribed antibiotics**				
	Bisoffi 2009	52.9%	1058	54.8%	1111
	Ansah 2010	26.6%	1719	22.3%	1723
1.5	**Patients still unwell at follow-up at day 4, or after**				
	Yeboah 2010	9.3%	1017	10.0%	2108
	Bisoffi 2009	5.6%	1024	5.5%	1071

2.0 SUBGROUPS

2.1	**Fever patients getting anti-malarials**				
2.1.1	<5 years				
	Yeboah-Antwi 2010*	27.5%	963	99.1%	2084
	Ansah 2010	80.0%	519	92.5%	550
2.1.2	5 + years				
	Ansah 2010	65.8%	1200	92.8%	1173
	Skarbinski 2009	40.9%	359	54.2%	310

*RRs calculated from these summary statistics may be different from the ones in the analysis. In the analysis, the review authors extracted RRs which had been adjusted for clustering and baseline imbalance

Appendix 9 Description of the UMSP trial (Chapter 4)

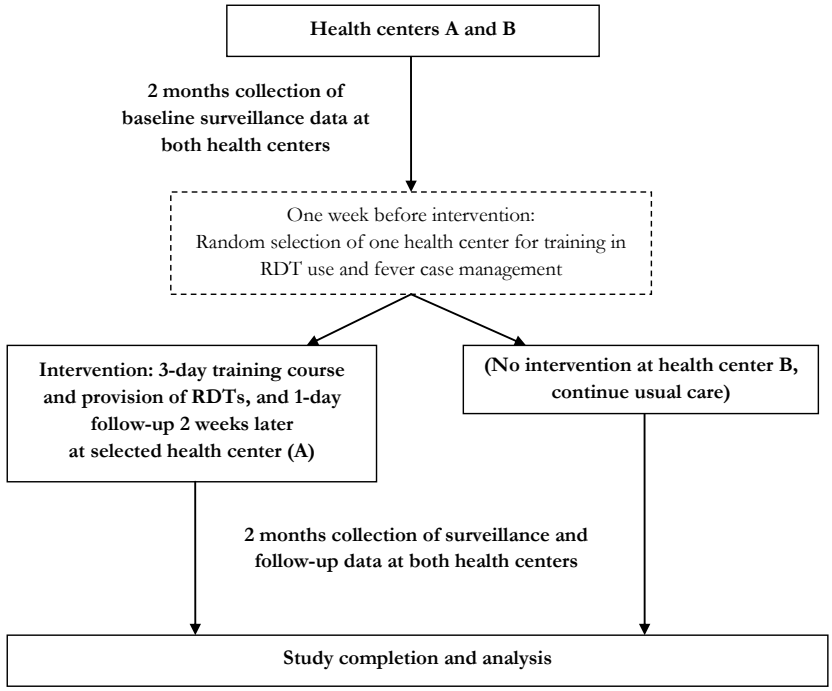
Settings

The trials were run in 3 districts representing high (Tororo), medium (Jinja) and low (Mubende) transmission settings

Design

The study was a cluster randomised control trial, with parallel baseline and follow-up surveys at each study site. It was conducted between 9th January 2008 to 25th June 2008, lasting 4 months in each facility (2 months of pre-test and 2 months of pos-test data collection).

Figure 28: Design of the trial conducted by the Uganda Malaria Surveillance Project in 2008



Source: Dorsey and Hopkins [81]

In each of the 3 zones, two health centres (A and B) were identified for the evaluation. The health centres included were at referral levels II and III (HC II and HC III), all of which without the capacity for microscopy. Fever treatment was presumptive prior to the study. The health centres (HC) were included if they were within 15 km radius of the UMSP sentinel site, had a workload of ≥ 500 patients per month and had at least 3 clinical staff.

Baseline Survey

The trial was preceded by 2 months of baseline survey during which surveillance data were collected on every outpatient at the selected HCs, with regard to their age, gender, symptoms, diagnosis, and treatment. Follow-up evaluations of patients' clinical outcomes were NOT conducted during the baseline survey; neither were parasitological confirmation and typing of malaria parasites done during the baseline survey.

Intervention

Following the baseline survey, one of the health centres in each district was randomly selected to receive training in RDT use and fever case management, and was subsequently provided with RDT kits

(Parachek®, HRP-2-based test)[17] and the RDT-based training/treatment manual[175]. The other health centre in each district continued with presumptive treatment based on IMCI algorithm.

All the staff from the intervention health centres received hands-on instruction on the preparation and interpretation of RDTs. The training focused on the criteria for selecting patients for RDT, hands-on preparation and interpretation of RDT results, and the management of fever patients with positive or negative RDT. Other topics included recognition of symptoms of severe conditions and the appropriate referral of patients with severe illness; counselling of patients on treatment; RDT storage and stock management.

The training lasted for 3 days; and was conducted by the same team of trainers, who were experienced in training and in the use of RDT. The hands-on instructions on the preparation and interpretation of the RDTs were based on the WHO pictorial training materials (as available at:

http://www.wpro.who.int/sites/rdt/using_rdts/RDT+Instructions+and+Training.htm)

At the intervention centres, blood samples were extracted from the participating patients for RDT; in addition, smears were made for reference microscopy and some blood was stored on filter paper for PCR to confirm parasitaemia and type malaria species. However, parasite confirmation and typing was NOT performed at the control health centres even during the intervention phase of the trials.

A follow up, one-day support training was offered to the clinicians in the intervention health centres at 2 weeks after the initial training, aimed at reinforcing the techniques for performing and interpreting RDTs. However, there was no interference with the clinicians' decisions—to test and/or treat patients according to the guideline.

Follow-up

Surveillance data collection continued for 2 more months after the introduction of RDT and the guideline at each intervention site. Data were regularly gathered on RDT performance at all the intervention sites; and on prescribing practices at all the study sites (both intervention and control facilities). At each of the intervention and control facilities, twenty five percent of patients were systematically recruited for follow-up survey to assess and document their clinical outcomes at day 5 post treatment initiation. The patients or their parents/caretakers were advised to return to the health centre for the follow-up evaluation 5 days after their initial contact with the health worker, or to seek appropriate care earlier if their conditions deteriorated. During the follow-up visits, participants were asked to describe their response to therapy received on day 0 encounters with the clinician, including any intervening care-seeking behaviour (e.g. additional health-care sought or medications taken following the initial day 0 encounters). A participant who did not turn up for the follow-up evaluation by day 6 was considered lost to follow-up.

Participants

All ambulatory patients presenting at the facilities were included in the study, except those who presented with severe conditions, or did not consent[18]. Patients were sequentially recruited into the study, as long as they satisfied the aforementioned criteria.

Sample size

Based on previous surveillance work in Uganda, it was estimated that approximately 1000 patients would be seen each month at each health centre[19] [81]; and that during the two months of baseline surveillance, 30%, 40% and 60% of patients would be treated for malaria at the health centers from the low, medium

[17] Manufactured by Orchid Biomedicals – Goa, India

[18] However, for the purposes of this report, analysis of effects was based on participants presenting with fever only.
[19] i.e. 2000 during baseline and 2000 after introducing the intervention

and high transmission intensity sites, respectively. Firstly, the trial aimed to test the hypothesis that the intervention of RDTs and training would lead to a significant reduction in the proportion of patients treated for malaria, compared to the non-intervention groups at each of the 3 malaria transmission intensity areas. Secondly, it aimed to test the hypothesis that the intervention would result in no significant difference in the proportion of patients with inadequate response to initial therapy, compared to the non-intervention group (i.e. non-inferiority analysis). The study was powered at 80% to detect a 10% reduction in the proportion of patients treated in the intervention arm at one-sided significance level of 5%; and to show non-inferiority between the study arms at two-sided significance level of 5%. It was assumed that in the non-intervention group 10% of patients would have an inadequate response to initial therapy, and that the proportion of patients with an inadequate response to initial therapy in the intervention group would be no higher than 15.3%. On the basis of these assumptions, the minimum number of observations required to show a significant reduction in the proportion of patients treated for malaria in the intervention arm was calculated to be 92. However, all consenting eligible patients attending the clinics during the study periods were recruited. The study planned to select 25% of all enrolled patients for evaluation of clinical outcomes at day 5 (i.e. every fourth of the expected 2000 patients during the post-intervention phase, giving an estimated 500 patients per study site). It was estimated that 10% of patients would refuse to participate and an additional 10% of patients would not return for their 5-day follow-up interview, leaving a sample size of 400 for each site for outcomes evaluation.

Quality control

Members of the study teams were trained on the study protocol prior to the onset of the study. During the study, case record forms and books were reviewed by the study coordinator and/or assistants for completeness and accuracy. The study coordinator met frequently with the site staff to ensure consistency in data collection. To ensure quality of reference microscopy (the gold standard test), each research slide was independently reviewed by two expert microscopists, and any discrepancies in slide readings were reviewed and resolved by a third microscopist. Before the beginning of the study at each site, positive and negative blood samples were obtained and standard aliquots frozen and stored in Kampala for quality control testing of RDTs[20]. Each lot of RDTs underwent quality control testing according to WHO guidelines.[21]

Data entry and characteristics of the trial database used in the analysis

Data from the survey case records were double-entered into EPI INFO 6.04d to verify their accuracy. A check program was written into the database to limit the entry of incorrect data and ensure entry of data into required fields. Two back-up files of the database were created after each data entry session.

The baseline and follow-up survey data were entered into two separate databases; and a third one where baseline and follow-up survey data sets were merged. This analysis used the merged database. It provides data on the total number of patients seen over the 4 months study period, their demographic profiles, the presenting symptoms, the diagnosis made and the type of treatment provided. In addition, it provides data on whether an RDT was performed on a patient, the RDT result, treatment given, and outcome of illness at day 5 post-treatment. Furthermore, it provides data on the malaria status of participants during the follow-up surveys, but at the intervention sites only[22].

---

[20] HRP2 levels remain stable in frozen blood for at least a year (81. Dorsey, G. and H. Hopkins, *Effectiveness and safety of training in fever case management incorporating rapid diagnostic tests (RDTs) for malaria at peripheral health centers in Uganda; draft version 1.0*, 2007, Uganda Malaria Surveillance Project (UMSP).

[21] World Health Organization. Malaria rapid diagnosis: Making it work. WHO informal consultations, 2003.

[22] Parasitological conformation was undertaken to assess the accuracy of RDT; and so was performed during the follow-up surveys at the intervention sites only

Appendix 10 Unit prices of items prescribed in trial

Category	Item	Prices (US$)*	Unit	Information Source
DIAGNOSTICS	RDTs	$ 0.6	Test	NMS
ANTIBIOTICS	Ampicillin	$ 0.01	Capsule	NMS
	Ampiclox	$ 0.03	Capsule	NMS
	Amoxycillin	$ 0.02	Capsule	NMS
	C/amphenical	$ 0.01	Capsule	NMS
	Ciprofloxacin	$ 0.03	Tablet	NMS
	Cotrimoxazole	$ 0.01	Tablet	NMS
ANTI-MALARIALS	AL	$ 0.084	Tablet	NMS
	CQ	$ 0.00	Tablet	NMS
	Quinine	$ 0.04	Tablet	NMS
	SP	$ 0.02	Tablet	NMS
ANALGESICS	Aspirin	$ 0.001	Tablet	NMS
	Indomethacin	$ 0.003	Tablet	NMS
	Ibuprofen	$ 0.003	Tablet	NMS
	Paracetamol	$ 0.003	Tablet	NMS
	Diclofenac	$ 0.046	Tablet	NMS

Information Source:
NMS=National Medical Stores catalogue and price indicator of 2008

*Exchange rate = US$ 1 = 2177.56 UGX (Uganda Shillings)

Appendix 11 Recommended quantities of items as per MOH guideline

	Age Groups				
MEDICATION	6 mo – 3 yrs	>3yrs – 5yrs	>5yrs - 7yrs	>7yrs–12 yrs	>12 yrs
RDT	1	1	1	1	1
AL	6	12	12	18	24
CQ	3	5	6	7	10
Quinine	3.75	7.5	7.5	15	30
SP	0.5	1	1.5	2	3
CTX (400+80 mg)*	2.5	5	5	10	20
Other Antibiotics**	5	10	10	20	40
Aspirin	2.25	4.5	4.5	9	18
Indomethacin	2.25	4.5	4.5	9	18
Ibuprofen	2.25	4.5	4.5	9	18
Paracetamol	2.25	4.5	4.5	9	18
Diclofenac	2.25	2.25	4.5	4.5	9

* CTX = co-trimoxazole

** In general, CTX was prescribed about 1.6 times more often than all the other antibiotics combined. For the purpose of this analysis, all antibiotics are grouped into 2 categories: co-trimoxazole and others. The analysis applies the average price for other listed antibiotics to all those that fall in the "other" category. (Lubell et al (2008) applies the average price of antibiotics to the total number of recipients [140]).

The analysis assumes that a unit (tablet/capsule) of each "other antibiotics" contains 250mg of the active agent; and that each of the "other antibiotics" were prescribed 6 hourly

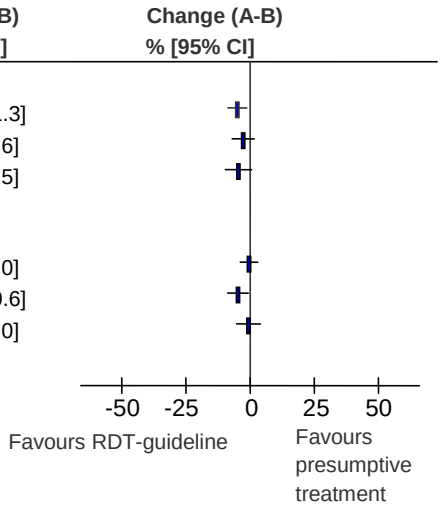
Appendix 12 Patients prescribed antimalarials before and after intervention by age and sites

	Age (Yrs)	**TORORO** (High transmission) Before %	(N)	After %	(N)	**Change* (%)**	**JINJA** (Medium transmission) Before %	(N)	After %	N	**Change* (%)**	**MUBENDE** (Low transmission) Before %	(N)	After %	(N)	**Change* (%)**
Intervention Sites	<5	91.0	(444)	83.3	(1360)	**-7.7**	92.0	(238)	55.3	(367)	**-36.7**	98.9	(183)	46.9	(160)	**-52.0**
	5+	94.5	(541)	52.4	(928)	**-42.1**	93.8	(768)	38.0	(706)	**-55.8**	99.5	(759)	26.2	(442)	**-73.3**
	All	92.9	(985)	70.8	(2288)	**-22.1**	93.3	(1006)	44.4	(1073)	**-48.9**	99.4	(942)	39.2	(602)	**-60.2**
Control Sites	<5	91.1	(629)	93.4	(934)	**2.3**	96.0	(272)	98.3	(354)	**2.3**	98.3	(119)	99.1	(214)	**0.8**
	5+	86.9	(817)	95.8	(975)	**8.9**	97.1	(896)	97.8	(786)	**0.7**	99.7	(652)	99.3	(734)	**-0.4**
	All	88.7	(1446)	94.6	(1909)	**5.9**	97.3	(1168)	98.3	(1140)	**1.0**	99.5	(771)	99.3	(948)	**-0.2**

*percentage change from baseline

Appendix 13 Patients prescribed antibiotics before and after introducing RDTs

Sites & settings	Before (B) %	Before (B) Total	After (A) %	After (A) Total	Change (A-B) % [95% CI]
Intervention sites					
High transmission	43.7	985	38.7	2288	-5.0 [-8.7, -1.3]
Medium transmission	50.7	1006	48.0	1075	-2.7 [-7.0, 1.6]
Low transmission	51.4	942	46.8	602	-4.6 [-9.7, 0.5]
Control sites					
High transmission	53.1	1446	52.7	1909	-0.4 [-3.8, 3.0]
Medium transmission	55.1	1168	50.4	1140	-4.7 [-8.8, -0.6]
Low transmission	39.5	771	38.9	948	-0.6 [-5.2, 4.0]



Appendix 14 Average treatment cost before and after intervention

Site	District	Setting	Average cost (2008 US$) Before (B)	After (A)	Change (A-B)	(95% CI)
Intervention	Tororo	High	1.47	1.49	0.02	(-0.21, 0.25)
	Jinja	Medium	1.86	1.53	-0.33	(-0.54, -0.12)
	Mubende	Low	2.04	1.54	-0.50	(-0.69, -0.31)
Control	Tororo	High	1.44	1.43	-0.02	(-0.19, 0.17)
	Jinja	Medium	1.97	1.74	-0.22	(-0.40, -0.06)
	Mubende	Low	2.07	1.91	-0.16	(-0.39, 0.07)

Appendix 15 Average cost at intervention HCs before and after introducing RDTs, by age group

			Average cost (2008 US$)	
District	**Setting**	**Age**	Before (B)	After (A)
Tororo	High	<5	0.52	1.13
		5+	2.25	2.02
Jinja	Medium	<5	0.54	1.01
		5+	2.27	1.80
Mubende	Low	<5	0.57	0.99
		5+	2.39	1.74

Appendix 16 Perceived clinical status of patients at day 5 of follow-up

		TORORO (High transmission)		**JINJA** (Medium transm)		**MUBENDE** (Low transmission)	
Age group	Status	Control	Intervention	Control	Intervention	Control	Intervention
< 5	Improved	93.8	92.4	97.8	97.3	97.4	100
	No change	0.5	0.3	0.0	0.0	2.6	0.0
	Worse	5.2	7.0	2.2	2.7	0.0	0.0
	Can't tell	0.5	0.3	0.0	0.0	0.0	0.0
	(N)	(211)	(342)	(93)	(74)	(39)	(33)
5+	Improved	83.6	88.2	94.8	95.8	94.6	96.3
	No change	1.1	2.2	0.6	0.0	0.0	0.0
	Worse	14.7	9.6	4.6	3.5	5.4	3.7
	Can't tell	0.6	0.0	0.0	0.7	0.0	0.0
	(N)	(177)	(229)	(174)	(142)	(148)	(108)

Appendix 17 Information sheet and consent form (Chapter 5 - HFA)

CONSENT FOR STRCUTURED OR INDEPTH INTERVIEWS WITH THE PRESCRIBER AND/OR STAFF INCHARGE OF THE HEALTH UNIT

(Administered by me (JOHN ODAGA))

Hello, it is a pleasure meeting you. My Name is JOHN ODAGA, a researcher from Uganda Martyrs University, and a student of Liverpool School of Tropical Medicine, UK. Thank you for sparing this time to participate in this interview.

I am carrying out a research on the management of fever ("temperature") in children in order to document the causes and treatment of fever in this Health Centre (HC). In addition, I would like to find out if the essential inputs and support required in managing a child with fever are adequate in this HC. Your HC was randomly selected to participate in this study.

I will be asking you questions about various activities and inputs which are required in the management of a child with fever. I would like to learn and document the process of managing a child with fever in this HC. For this reason, I will ask to observe 6 consultations of children with fever, with the most senior health worker. I understand that the District Health Office recently supplied this HC with stocks of RDTs and ACTs. Feel free to use them as you deem fit. In addition, we will ask to briefly interview the care takers of the children selected for the observation as they exit the facility. This will entail reviewing the prescriptions received by the children here today. Further, I will ask to go through your stock cards and to assess the consultation room and store for selected equipment.

No patient names will be reviewed, recorded, or shared. Data will be summarised in general terms without specifying the participating facilities or individuals. Neither your name nor the name of this HC will be provided, and any reports that use this facility's data will only present information in aggregate form so that the facility cannot be identified. Your name and all information that you give me will be kept strictly confidential.

The information gathered may not be immediately and directly beneficial to you, or this HC. The information will only be used for the purpose of the research. The information may be used by the MOH for planning service improvement or for further studies of health services.

I have come along with a research assistant who will assist me in observing the consultations and in conducting exit surveys with the caretakers, and to review supplies and clinic attendance records. Please, I would like to request you to explain the purpose of this research to the care takers of selected patients; and that an observer will be sitting in the consultation room to observe the consultation process.

You or/and patients' care takers are free to refuse to participate in the study, or to withdraw at any stage, and this will not affect our opinion about you or the health centre. If you have further queries in future, you are free to contact me on 0772 619 450. Otherwise, you are free to ask any questions right now.

Declaration of Client

I have understood the descriptions above about the study. I have had time to ask questions and they have been answered to my satisfaction. I agree to take part in the study.

Printed Name and Title of Informant
…….

Date ……. Signature …….

Printed Name of Person obtaining consent
…….

Date …….
Signature…….

MODULE 1: CLINICAL OBSERVATION OF SIX CONSECUTIVE FEBRILE CHILDREN

Facility Code: | **Interviewer Code:**

Date of Observation **dd** **mm** **Year**

READ CONSENT FORM TO HEALTH WORKER. READ CONSENT FROM TO THE CHOSEN CARETAKERS *BEFORE* THEY ENTER THE CONSULTATION ROOM.

OBSERVE SIX CONSECUTIVE ELIGIBLE CLINICAL CASES.

ELIGIBLE CASES ARE ARE SICK CHILDREN, 1-59 MONTHS OF AGE, WHO HAVE BEEN BROUGHT TO THE CLINIC WITH COMPLAINT OF **FEVER OR A HISTORY OF FEVER.**
THEY MAY OR MAY NOT HAVE OTHER COMPLIANTS

THERE IS A SEPARATE COLUMN FOR EACH OF THE SIX CASES OBSERVED. WHERE RELEVANT, CIRCLE **Y**ES, **N**O, OR **N**OT **A**PPLICABLE FOR EACH QUESTION.

	QUESTIONS	CODING CLASSIFICATION					
100	**GENERAL INFORMATION AND PRESENTING COMPLAINTS**	**CASE 1**	**CASE 2**	**CASE 3**	**CASE 4**	**CASE 5**	**CASE 6**
100A	RECORD THE EXACT TIME THAT THE CARETAKER ENTERS THE EXAMINATION ROON	TIME:	TIME:	TIME:	TIME:	TIME:	TIME:
101	A. WHAT TYPE OF HEALTH WORKER EXAMINED THE CHILD? Write the letter code corresponding to the appropriate title below: **A** = Clinical officer; **B** = Reg. Nurse/Mid wife; **C** = Enrolled Nurse/Midwife; **D** = Nursing Assistant; **E** = Other (specify)						
102	AGE OF CHILD (IN COMPLETED MONTHS - 1 TO 59)-- ASK THE CARE TAKER IF THE CHILD'S AGE IS NOT RECORDED ON THE MEDICAL FORM						
103	A. WEIGHT OF CHILD (IN KG) -- RECORD **"NA"** if not indicated or taken						
104	REASON FOR VISIT (Circle the letter code(s) corresponding to ALL the condition/s that apply) *(SHOULD ONLY BE FOR CASES WITH **FEVER/MALARIA, WITH OR WITHOUT** OTHER CONDITIONS)* **A** = FEVER/MALARIA ONLY; **B** = FEVER + COUGH/DIFFICULT IN BREATHING; **C** = FEVER + DIARRHOEA; **D** = FEVER+COUGH+DIARRHOEA; **E** = FEVER+ANY OTHER CONDITION	A B C D E	A B C D E	A B C D E	A B C D E	A B C D E	A B C D E

	PATIENT EVALUATION:--HISTORY TAKING	CASE 1			CASE 2			CASE 3			CASE 4			CASE 5			CASE 6		
	QUESTIONS																		
105	DOES THE HEALTH WORKER.																		
	A. ASK ABOUT FEVER OR HISTORY OF FEVER (**NA** if caretaker volunteers this information while in the consultation room before HW asks)?	Y	N	NA	Y	N	NA	Y	N	NA	Y	N	NA	Y	N	NA	Y	N	NA
	B. ASK ABOUT COUGH OR DIFFICULT BREATHING OR HISTORY (**NA** if caretaker volunteers this information while in the consultation room before HW asks)?	Y	N	NA	Y	N	NA	Y	N	NA	Y	N	NA	Y	N	NA	Y	N	NA
	C. ASK ABOUT DIARRHEA OR HISTORY OF DIARRHEA (**NA** if caretaker volunteers this information while in the consultation room before HW asks)?	Y	N	NA	Y	N	NA	Y	N	NA	Y	N	NA	Y	N	NA	Y	N	NA
	D. ASK ABOUT THE ABILITY TO FEED OR BREASTFEED (**NA** if caretaker volunteers this information while in the consultation room before HW asks; or if child >2 yrs)?	Y	N	NA	Y	N	NA	Y	N	NA	Y	N	NA	Y	N	NA	Y	N	NA
	E. ASK WHETHER THE CHILD VOMITS EVERYTHING (**NA** if caretaker volunteers this information while in the consultation room before HW asks)?	Y	N	NA	N	N	NA	N	N	NA	N	N	NA	N	N	NA	N	N	NA
	F. ASK ABOUT THE PRESENCE OF CONVULSIONS (**NA** if caretaker volunteers this information while in the consultation room before HW asks)?	Y	N	NA	Y	N	NA	Y	N	NA	Y	N	NA	Y	N	NA	Y	N	NA
	PATIENT EVALUATION:--PHYSICAL EXAMINARION	**CASE 1**			**CASE 2**			**CASE 3**			**CASE 4**			**CASE 5**			**CASE 6**		
	QUESTIONS																		
106	DOES THE HEALTH WORKER																		
	A. CHECK THE EYES/CONJUCTIVE/PALMS/FINGER NAILS, SOLES FOR PALLOR AND JAUNDICE?	Y	N		Y	N		Y	N		Y	N		Y	N		Y	N	
	B. CHECK THE THROAT FOR REDNESS	Y	N		Y	N		Y	N		Y	N		Y	N		Y	N	
	C. CHECK THE EARS FOR DISCHARGE/SORES	Y	N		Y	N		Y	N		Y	N		Y	N		Y	N	
	E. COUNT RESPIRATORY RATE WITH TIMING DEVICE? (**NA** if no cough or difficult breathing)	Y	N	NA	Y	N	NA	Y	N	NA	Y	N	NA	Y	N	NA	Y	N	NA
107	DOES THE HEALTH WORKER REQUEST FOR/PERFRM AN RDT OR MICROSCOPY TO RULEOUT MALARIA?; IF 'NO' JUMP TO SECTION 109	Y	N		Y	N		Y	N		Y	N		Y	N		Y	N	
108	TEST RESULT: POS = POSITIVE; NEG = NEGATIVE; INT = INDETERMINATE	POS	NEG	INT	POS	NEG	INT	POS	NEG	INT	POS	NEG	INT	POS	NEG	INT	POS	NEG	INT
109	DID THE HEALTH WORKER COMMUNICATE THE RESULT OF THE RDT TEST TO THE PATIENT/CARETAKER	Y	N		Y	N		Y	N		Y	N		Y	N		Y	N	

	CLASSIFICATION AND TREATMENT OF PATIENT QUESTIONS	CASE 1			CASE 2			CASE 3			CASE 4			CASE 5			CASE 6		
110	DOES THE HEALTH WORKER CLASSIFY THE CHILD AS HAVING (Fill in the appropriate Code)																		
	A. FEVER OR (uncomplicated) MALARIA ONLY?																		
	B. FEVER/MALARIA + OTHER CONDITIONS (e.g. ACUTE RESPIRATORY INFECTIONS)																		
	C. ACUTE RESPIRATORY INFECTIONS ONLY (ARI)--including Pneumonia																		
	D. OTHER CONDITIONS, EXCLUDING MALARIA AND/OR ARI (SPECIFY)																		
	[NB. IF THE HEALTH WORKER CLASSIFYS THE CHILD AS HAVING SEVERE DISEASE (SEVERE PNEUMONIA, SEVERE MALARIA, OR SEVERE FEBRILE ILLNESS) EXCLUDE HIM/HER FROM THE OBSERVATION]																		
	TREATMENT (PRECRIPTION AND COUNSELLING)																		
111	DOES THE HEALTH WORKER PRESCRIBE																		
	A. ACT FIXED COMBINATION: COARTEM? Colour Codes: Yellow/Blue/Brown/Green	Y	N	NA	Y	N	NA	Y	N	NA	Y	N	NA	Y	N	NA	Y	N	NA
	B. ACT FIXED COMBINATION: ARTESUNATE-AMODIAQUINE?	Y	N	NA	Y	N	NA	Y	N	NA	Y	N	NA	Y	N	NA	Y	N	NA
	C. OTHER ANTI-MALARIA DRUGS	Y	N	NA	Y	N	NA	Y	N	NA	Y	N	NA	Y	N	NA	Y	N	NA
	IF YES, PLEASE SPECIFY THE APPROPRIATE DRUGS IN THE	.....			.....			.....			.....			.....			.....		
	SPACES PROVIDED (examples are given below)	.....			.....			.....			.....			.....			.....		
	[Examples; Artemether monotherapy, Artesunate monotherapy, Choloroquine, SP, Amodiaquine, Glucodexine, etc]	.....			.....			.....			.....			.....			.....		
	C. COTRIMOXAZOLE OR AMOXYCILLIN ?	Y	N	NA	Y	N	NA	Y	N	NA	Y	N	NA	Y	N	NA	Y	N	NA
	E. CIPROFLOXACIN?	Y	N	NA	Y	N	NA	Y	N	NA	Y	N	NA	Y	N	NA	Y	N	NA
	C. ANTIBIOTICS	Y	N	NA	Y	N	NA	Y	N	NA	Y	N	NA	Y	N	NA	Y	N	NA
	IF YES, PLEASE SPECIFY THE APPROPRIATE DRUGS IN THE	.....			.....			.....			.....			.....			.....		
	SPACES PROVIDED (examples are given below)	.....			.....			.....			.....			.....			.....		
		.....			.....			.....			.....			.....			.....		
		.....			.....			.....			.....			.....			.....		
	D. ORS (or IV fluids - *only* in case of severe dehydration)?	Y	N	NA	Y	N	NA	Y	N	NA	Y	N	NA	Y	N	NA	Y	N	NA
112	DOES HEALTH WORKER EXPLAIN TO CARETAKER HOW TO GIVE A. ACT?	Y	N	NA	Y	N	NA	Y	N	NA	Y	N	NA	Y	N	NA	Y	N	NA
	B. OTHER ANTI-MALARIAL DRUGS?	Y	N	NA	Y	N	NA	Y	N	NA	Y	N	NA	Y	N	NA	Y	N	NA
	C. ANTIBIOTICS?	Y	N	NA	Y	N	NA	Y	N	NA	Y	N	NA	Y	N	NA	Y	N	NA
113	RECORD THE EXACT TIME THAT THE CONSULTATION ENDS	TIME:			TIME:			TIME:			TIME:			TIME:			TIME:		

Supervisor Recode for HW performance - treatment: Does classification (Q.110) match the medication prescibed (Q.111)?	**CASE 1**		**CASE 2**		**CASE 3**		**CASE 4**		**CASE 5**		**CASE 6**	
A. malaria or fever / ACT	Match	Not match	Match	Not match	Match	Not match	Match	Not match	Match	ot match	Match	Not match
B. malaria or fever + other conditions / ACT + antibiotics	Match	Not match	Match	Not match	Match	Not match	Match	Not match	Match	Not match	Match	Not match
C. ARI / antibiotics / no antibiotics	Match	Not match	Match	Not match	Match	Not match	Match	Not match	Match	Not match	Match	Not match
D. diarrhea **without** blood / ORS but no antibiotic	Match	Not match	Match	Not match	Match	Not match	Match	Not match	Match	Not match	Match	Not match
E. diarrhea **with** blood / Ciprofloxacin or other antibiotics	Match	Not match	Match	Not match	Match	Not match	Match	Not match	Match	Not match	Match	Not match
F. **No** pneumonia, fever/malaria or diarrhea / no antibiotic(unless indicated for other than pneumonia or dysentery) or no antimalarial	Match	Not match	Match	Not match	Match	Not match	Match	Not match	Match	Not match	Match	Not match
HW performance - treatment (All match (A-F)?)	Y	N	Y	N	Y	N	Y	N	Y	N	Y	N
# Y (out of 6 cases) _____												
NOTE ANY QUALITATIVE OBSERVATIONS HERE:												

MODULE 2: OBSERVATION OF RDT PERFORMANCE ON SIX CONSECUTIVE CHILDREN WITH FEVER

Facility Code: | | | | **Interviewer Code:** | | | |

Date of Observation | | | | **dd** **mm** **Year**

		CODING CLASSIFICATION																	
	QUESTIONS	**CASE 1**			**CASE 2**			**CASE 3**			**CASE 4**			**CASE 5**			**CASE 6**		
200	RECORD THE EXACT TIME THAT THE CARETAKER ENTERS THE TEST ROOM (IF TEST IS PERFORMED BY ANOTHER HEALTH WORKER IN A SEPARATE ROOM); OR THE TIME WHEN THE HEALTH WORKER STARTS TO PREPARE FOR THE TEST	TIME:			TIME:			TIME:			TIME:			TIME:			TIME:		
201	WAS THE RDT PERFORMED BY THE CLERKING HEALTH WORKER OR COMPLEMENTARY HEALTH WORKER	Y	N		Y	N		Y	N		Y	N		Y	N		Y	N	
202	WHAT TYPE OF HEALTH WORKER PERFORMED THE RDT? Write the letter code corresponding to the appropriate title below. Please, answer this question even if the Health worker concerned is the same one that clerked the patient **A** = Clinical officer; **B** = Reg. Nurse/Mid wife; **C** = Enrolled Nurse/Midwife; **D** = Nursing Assistant; **E** = Other (specify)																		
203	DOES THE HEALTH WORKER.																		
	A. CHECK THE EXPIRY DATE ON THE PACK	Y	N	NA	Y	N	NA	Y	N	NA	Y	N	NA	Y	N	NA	Y	N	NA
	B. PUT ON NEW GLOVES TO PERFORM RDT ON THIS PATIENT	Y	N	NA	Y	N	NA	Y	N	NA	Y	N	NA	Y	N	NA	Y	N	NA
	C. WRITE THE PATIENT'S NAME ON THE CASSETTE	Y	N	NA	N	N	NA	N	N	NA	N	N	NA	N	N	NA	N	N	NA
	D. CLEAN THE PATIENTS FINGER WITH A SPIRIT SWAB	Y	N	NA	Y	N	NA	Y	N	NA	Y	N	NA	Y	N	NA	Y	N	NA
	E. ALLOW THE PATENT'S FINGER TO DRY BEFORE PRICKING	Y	N	NA	Y	N	NA	Y	N	NA	Y	N	NA	Y	N	NA	Y	N	NA
	F. USE A LOOP TO COLLECT THE DROP OF BLOOD ON THE PATIENT'S FINGER	Y	N	NA	Y	N	NA	Y	N	NA	Y	N	NA	Y	N	NA	Y	N	NA
	G. PUT THE DROP OF BLOOD INTO THE SQUARE HOLE (AT POSITION A)	Y	N	NA	Y	N	NA	Y	N	NA	Y	N	NA	Y	N	NA	Y	N	NA
	H. PUT SIX (6) DROPS OF BUFFER INTO THE ROUND HOLE AT POSITION B	Y	N	NA	Y	N	NA	Y	N	NA	Y	N	NA	Y	N	NA	Y	N	NA
	I. NOTE DOWN THE TIME (IN MINS) WHEN THE HEALTH WORKER ADDS BUFFER ON TO THE RDT																		
	J. NOTE DOWN THE TIME WHEN THE HEALTH WORKER READS AND RECORDS THE RESULT OF RDT																		
	K. WAIT 15 MINUTES AFTER ADDING BUFFER (CAN BE RECORDED LATER)	Y	N	NA	Y	N	NA	Y	N	NA	Y	N	NA	Y	N	NA	Y	N	NA
204	RDT TEST RESULT: POS = POSITIVE; NEG = NEGATIVE; INT = INDETERMINATE	POS	NEG	INT	POS	NEG	INT	POS	NEG	INT	POS	NEG	INT	POS	NEG	INT	POS	NEG	INT
205	DID THE HEALTH WORKER COMMUNICATE THE RESULT OF THE RDT TEST TO THE PATIENT/CARETAKER	Y	N		Y	N		Y	N		Y	N		Y	N		Y	N	
	NOTE ANY QUALITATIVE OBSERVATIONS HERE:																		

MODULE 2: EXIT INTERVIEW (CARETAKERS OF SIX OBSERVED SICK CHILDREN)

Facility Code: | Interviewer Code: | Date: dd mm yy

OBTAIN INFORMED CONSENT FROM EACH CARETAKER IF THE SUPERVISOR HAS NOT ALREADY DONE SO.

NO.	QUESTIONS	CODING CLASSIFICATION (PUT CASE CODE AT TOP OF EACH COLUMN)					
		CASE 1	CASE 2	CASE 3	CASE 4	CASE 5	CASE 6
300	What illness(es) did the health worker tell you your child had? *[Refer to the Child's card and record the diagnosis given; otherwise list down the complaints recorded]*						
301	Did the Health worker perform a blood test on your child today? ***[Skip to 204 if the answer is NO]***	Y N	Y N	Y N	Y N	Y N	Y N
302	Did the Health worker explain to you what the test was for?	Y N	Y N	Y N	Y N	Y N	Y N
303	Did the health worker explain the result of the test to you?	Y N	Y N	Y N	Y N	Y N	Y N
304	Did the health worker give or prescribe any medicines for yor child today?	YES…… 1	YES…… 1	YES…… 1	YES…… 1	YES…… 1	YES…… 1
		NO (end interview) 2	NO (end interview) 2	NO (end interview) 2	NO (end interview) 2	NO (end interview) 2	NO (end interview) 2
		DON'T KNOW 3	DON'T KNOW 3	DON'T KNOW 3	DON'T KNOW 3	DON'T KNOW 3	DON'T KNOW 3
305	Did the health worker explain why he prescribed for you these particular medicines; or why he did not prescribe for you any medicines (or some medicines)?	Y N	Y N	Y N	Y N	Y N	Y N
306	Did you receive all the medications prescribed?	Yes, I Received all 1	Yes, I Received all 1	Yes, I Received all 1	Yes, I Received all 1	Yes, I Received all 1	Yes, I Received all 1
		No, I Received some 2	No, I Received some 2	No, I Received some 2	No, I Received some 2	No, I Received some 2	No, I Received some 2
		No, I did not receive any 3	No, I did not receive any 3	No, I did not receive any 3	No, I did not receive any 3	No, I did not receive any 3	No, I did not receive any 3
		I don't know if I received all 4	I don't know if I received all 4	I don't know if I received all 4	I don't know if I received all 4	I don't know if I received all 4	I don't know if I received all 4

NO.	QUESTIONS		CASE 1				CASE 2				CASE 3				CASE 4				CASE 5				CASE 6			
307	Can you please show me the medications or prescriptions given to you by the health worker?																									
	ASK THE MOTHER TO SHOW YOU EACH MEDICINE OR PRESCRIPTION GIVEN TO HER. THEN WRITE DOWN THE NAME OF EACH MEDICINE BELOW (UNDER "MEDICINE 1," "MED. 2," AND "MEDICINE 3.", ETC) ASK HER ABOUT THE AMOUNT TO BE GIVEN EACH TIME, THE NUMBER OF TIMES A DAY TO GIVE IT, AND THE NUMBER OF DAYS IT IS TO BE GIVEN.																									
			Med name:				Med name:				Med name:				Med name:				Med name:				Med name:			
01	WRITE NAME OF MEDICATION 1				Agreement				Agreement				Agreement				Agreement				Agreement				Agreement	
				MF 5	Yes	No		MF 5	Yes	No		MF 5	Yes	No		MF 5	Yes	No		MF 5	Yes	No		MF 5	Yes	No
	a. How much will you give each time?	MEDICATION 1	Amount:				Amount:				Amount:				Amount:				Amount:				Amount:			
	b. How many times a day will you give it?		#times/day:				#times/day:				#times/day:				#times/day:				#times/day:				#times/day:			
	c. For how many days will you give it?		#days:				#days:				#days:				#days:				#days:				#days:			
			Med name:				Med name:				Med name:				Med name:				Med name:				Med name:			
02	WRITE NAME OF MEDICATION 2				Agreement				Agreement				Agreement				Agreement				Agreement				Agreement	
				MF 5	Yes	No		MF 5	Yes	No		MF 5	Yes	No		MF 5	Yes	No		MF 5	Yes	No		MF 5	Yes	No
	a. How much will you give each time?	MEDICATION 2	Amount:				Amount:				Amount:				Amount:				Amount:				Amount:			
	b. How many times a day will you give it?		#times/day:				#times/day:				#times/day:				#times/day:				#times/day:				#times/day:			
	c. For how many days will you give it?		#days:				#days:				#days:				#days:				#days:				#days:			
			Med name:				Med name:				Med name:				Med name:				Med name:				Med name:			
03	WRITE NAME OF MEDICATION 3				Agreement				Agreement				Agreement				Agreement				Agreement				Agreement	
				MF 5	Yes	No		MF 5	Yes	No		MF 5	Yes	No		MF 5	Yes	No		MF 5	Yes	No		MF 5	Yes	No
	a. How much will you give each time?	MEDICATION 3	Amount:				Amount:				Amount:				Amount:				Amount:				Amount:			
	b. How many times a day will you give it?		#times/day:				#times/day:				#times/day:				#times/day:				#times/day:				#times/day:			
	c. For how many days will you give it?		#days:				#days:				#days:				#days:				#days:				#days:			
	Thank you for participating. We will use this information to help improve health services in this area.																									

	Supervisor Recode for HW performance - counseling		CASE 1		CASE 2		CASE 3		CASE 4		CASE 5		CASE 6	
	Is the caretaker's description of medication dose, frequency, and duration correct (Q.307)?	MED1	Correct	Not Corr.	Correct	Not Corr.	Correct	Not Corr.	Correct	Not Corr.	Correct	Not Corr.	Correct	Not Corr.
		MED2	Correct	Not Corr.	Correct	Not Corr.	Correct	Not Corr.	Correct	Not Corr.	Correct	Not Corr.	Correct	Not Corr.
		MED3	Correct	Not Corr.	Correct	Not Corr.	Correct	Not Corr.	Correct	Not Corr.	Correct	Not Corr.	Correct	Not Corr.
	HW performance - counselling: Knowledge correct for all medications? (Y/N)		Y	N	Y	N	Y	N	Y	N	Y	N	Y	N
	# Y out of 6 observations ____													
	NOTE ANY QUALITATIVE OBSERVATIONS HERE:													

MODULE 3: HEALTH FACILITY CHECKLIST (INFRASTRUCTURE, EQUIPMENT, SUPPLIES, DRUGS)

Facility Code: ____ **Interviewer Code:** ____ **Date:** ____ (dd) ____ (mm) ____ (yy)

OBTAIN INFORMED CONSENT

NO.	QUESTIONS	CODING CLASSIFICATION		GO TO
309	Can you please show me where children are seen for treatment INSPECT FOR AUDITORY AND VISUAL PRIVACY. MARK AS "BOTH" IF THERE IS A DOOR THAT CAN CLOSE MARK AS "VISUAL" IF THERE IS A DRAPE OR CURTAIN	VISUAL AND AUDITORY PRIVACY **VISUAL BUT NOT AUDITORY** PRIVACY VISUAL NOR AUDITORY PRIVACY	....	1 2 3	

IN THE CHILD CONSULTATION AREA, CHECK WHETHER EACH OF THE ITEMS BELOW IS EITHER IN THE ROOM WHERE THE SERVICE IS GIVEN OR IN AN ADJACENT ROOM.

	ITEMS FOR SICK CHILD CONSULTATIONS	(a) AVAILABILITY: OBSERVED	REPORTED, NOT SEEN	NOT AVAILABLE	DON'T KNOW	(b) FUNCTIONING: YES	NO	DON'T KNOW
01	Infant scale that is accessible	1 --> b	2 --> b	3	9	1	2	9
02	Adult (standing) scale that is accessible	1 --> b	2 --> b	3	9	1	2	9
03	Timer or Clock/watch with second hand	1 --> b	2 --> b	3	9	1	2	9
04	Thermometer	1 --> b	2 --> b	3	9	1	2	9
05	Refrigerator or cold box for storing vaccines	1 --> b	2 --> b	3	9	1	2	9
06	Cup and spoon for oral rehydration	1	2	3	9	1	2	9
07	Jar or pitcher for oral rehydration solution (ORS)	1	2	3	9	1	2	9

	INSPECT THE CONSULTATION AREA FOR THE PRESENCE OF THE FOLLOWING (WALL) CHARTS	OBSERVED AND IN CONS. AREA	OBSERVED BUT NOT IN CONS. AREA	REPORTED, NOT SEEN	NOT AVAILABLE	DON'T KNOW
01	RDT-based Treatment Algorithm	1	2	3	4	9
02	IMCI (sick child) Treatment Algorithm	1	2	3	4	9

	ASK TO SEE THE FOLLOWING GUIDELINES	OBSERVED AND IN CONS. AREA	OBSERVED BUT NOT IN CONS. AREA	REPORTED, NOT SEEN	NOT AVAILABLE	DON'T KNOW
01	MOH RDT/ACT USER'S MANUAL	1	2	3	4	9
02	Wall charts for RDT based treatment with ACT	1	2	3	4	9

313 ASK TO OBSERVE THE ROOM AND FACILITIES WHERE RDT SUPPLIES ARE KEPT

		OBSERVED				NOT OBSRVED
01	Describe the type of storage facility where the RDT kits are kept (e.g. in an ordinary cardboard box, in a cool on a shelf)	CARD BOARD BOX	COOL BOX	OPEN AIR (e.g. on a shelf, table, etc)	OTHER (Describe)	
		1	2	3	4	9
02	Roofing material	IRON SHEETS	GRASS	OTHER		NOT OBSRVED
		1	2	3	4	9
03	Presence of a ceiling	YES	NO	DON'T KNOW	9	NOT OBSRVED
		1	2	9		9
04	Check if the temperature of the room is recorded	RECORDED TO DAY				RECORDED OVER THE LAST 3 MONTHS
		1 YES (Records the highest temprature noted, if recorded more than once .....)				1 YES (Record the highest temperature) .....
		2 NO				2 NO
05	Amount of light in the storage room	DIM LIGHT	BRIGHT LIGHT			NOT OBSERVED
		1	2			9

--> go to b (to indicate whether functioning or not)

ASK TO SEE THE FOLLOWING DRUGS AND SUPPLIES. IF THE ITEM IS LOCATED IN A DIFFERENT PART OF THE FACILITY, GO THERE TO OBSERVE IT.
IF YOU ARE UNABLE TO SEE AN ITEM, ASK IF IT IS AVAILABLE AND THE EXPIRATION DATES HAVE NOT PASSED. FOR EACH ITEM, CIRCLE THE APPROPRIATE CODE.

314	INDICATOR DRUGS AND RDT KITS		(a) EXACT QUANTITY AVAILABLE TODAY (in single units, e.g tablets, capsules, vials)		(b) TOTAL NUMBER OF DAYS OUT OF STCOK IN THE LAST 6 MONTHS		(c) NOT OBSERVED/COUNTED
01	COARTEM						
	A) Yellow blister Packs	1	Number of doses	1	Number of days O/S		
		2	Not Counted	2	Not Counted		
	B) Blue Blister Packs	1	Number of doses	1	Number of days O/S		
		2	Not Counted	2	Not Counted		
	C) Brown Blister Packs	1	Number of doses	1	Number of days O/S		
		2	Not Counted	2	Not Counted		
	D) Green Blister Packs	1	Number of doses	1	Number of days O/S		
		2	Not Counted	2	Not Counted		
02	AMODIAQUINE-ARTESUNATE COMBINATION	1	Number of doses	1	Number of days O/S		
		2	Not Counted	2	Not Counted		
03	OTHER ANTI-MALARIA DMEDICINES						
	A) Other anti-malaria drugs(Tab Quinine)	1	Number of Tablets	1	Number of days O/S		
		2	Not Counted	2	Not Counted		
	B) Other anti-malaria drugs (Name)	1	Number of Tablets	1	Number of days O/S		
		2	Not Counted	2	Not Counted		
	C) Other anti-malaria drugs (Name)	1	Number of Tablets	1	Number of days O/S		
		2	Not Counted	2	Not Counted		
04	RDT kits	1	Number of kits	1	Number of days O/S		
		2	Not Counted	2	Not Counted		
05	COTRIMOXAZOLE FOR ARI	1	Number of Tablets	1	Number of days O/S		
	A) For Childrfen (Junior Tablets)						
		2	Not Counted	2	Not Counted		
	B) For Children (Syrups)	1	Number of Vials	1	Number of days O/S		
		2	Not Counted	2	Not Counted		
	C) For Adults (Tablets)	1	Number of Tablets	1	Number of days O/S		
		2	Not Counted	2	Not Counted		
06	ORS PACKETS FOR DIAORRHEA	1	Number of kits	1	Number of days O/S		
		2	Not Counted	2	Not Counted		

316	**ASK TO SEE THE RECORD SYSTEM FOR ORDERING AND ACCEPTING DELIVERY OF DRUGS AND SUPPLIES. IF YOU ARE NOT ABLE TO SEE THE RECORDS THEN ASK FOR THE FOLLOWING INFORMATION. REVIEW THE FIRST LINE ANTI-MALARIAL (ACT)**								
	REVIEW ENTRIES ABOUT ACTs								
			OBSERVED			**REPORTEDLY AVAILABLE, NOT OBSERVED**			
01	Review the date of the most recent order of anti-malarial / ACT and tell me the date please.	DATE IS IN LAST 3 MO. (1)	DATE IS MORE THAN 3 MO. (2)	NO DATE (3)		DATE IS IN LAST 3 MO. (4)	DATE IS MORE THAN (5)	NO DATE (6)	NO RECORD AVAILABLE (7)
02	Review the quantity of the most recent delivery . Is this amount the same as the quantity ordered?	DELIVERY AGREES WITH ORDER (1)	DELIVERY DOES NOT AGREE WITH ORDER (2)	QUANTITY IN ORDER MISSING (3)		DELIVERY AGREES WITH ORDER (4)	DELIVERY DOES NOT AGREE WITH ORDER (5)	QUANTITY IN ORDER MISSING (6)	NO RECORD AVAILABLE (7)
03	Does the Balance as recorded in the Records Agree with the quantity in the Stores/Pharmacy (CARRY OUT A HAND COUNT IF POSSIBLE OR ASK YOUR INFORMANT TO DO SO)	YES AGREE (1)	NO DOES NOT AGREE (2)			YES AGREE (4)	NO DOES NOT AGREE (5)		NO RECORD AVAILABLE (7)
04	Does the Balance as recorded in the Bin Card Agree with the quantity in the Stores/Pharmacy (CARRY OUT A HAND COUNT IF POSSIBLE OR ASK YOUR INFORMANT TO DO SO)	YES AGREE (1)	NO DOES NOT AGREE (2)			YES AGREE (4)	NO DOES NOT AGREE (5)		NO BIN CARD AVAILABLE (7)
05	Have any of the First Line Anti-Malarials / ACT passed their expiration date?	YES (1)	NO (2)	NO DATE VISIBLE (3)		YES (4)	NO (5)	NO DATE VISIBLE (6)	
06	Does this Health Facility have a plan to dispose of expired Drugs?	YES	NO						
	IF THE RESPONSE IS YES THEN ASK What is that plan?	APPROPRIATE PLAN EXISTS (1)	APPROPRIATE PLAN DOES NOT EXIST (2)	DID NOT KNOW (3)					
07	Please show me where or how expired drugs are destroyed?	APPROPRIATE PLAN IN USE (1)	NO APPROPRIATE PLAN IN USE (2)	COULD NOT OBSERVE (3)					

MODULE 4: HEALTH WORKER INTERVIEW & RECORD REVIEW

Facility Code: ____ **Interviewer Code:** ____

SPEAK TO THE MOST EXPERIENCED HEALTH WORKER INVOLVED IN MANAGEMENT OF CURATIVE CHILD HEALTH SERVICES.
IT IS BEST TO APPLY THIS FORM AFTER PATIENT SESSIONS HAVE FINISHED.
OBTAIN INFORMED CONSENT, IF YOU HAVE NOT ALREADY DONE SO.

NO.	QUESTIONS	CODING CLASSIFICATION	GO TO
401	For each of the following services, please tell me whether the service is offered by your facility, and if so, how many days per month the service is provided *at the facility, or as outreach services.* FOR THE PURPOSES OF THIS QUESTION, A MONTH IS EQUIVALENT TO FOUR WORK WEEKS.		
01	Consultation or curative services for sick children IF NONE, WRITE "00" IF ALL **WEEKDAYS**, WRITE "20" IF **ALL DAYS including weekends**, WRITE "30" IF ONE TIME PER WEEK, WRITE "4"	A. # OF DAYS PER MONTH IN FACILITY ____ B. # DAYS PER MONTH IN OUTREACH LOCATIONS ____	
02	Routine immunizations for children IF NONE, WRITE "00" IF ALL **WEEKDAYS**, WRITE "20" IF **ALL DAYS including weekends**, WRITE "30" IF ONE TIME PER WEEK, WRITE "4"	A. # OF DAYS PER MONTH IN FACILITY ____ B. # DAYS PER MONTH IN OUTREACH LOCATIONS ____	
03	Growth monitoring & promotion - where a healthy child is routinely weighed, has weight charted on growth chart, feeding advice given IF NONE, WRITE "00" IF ALL **WEEKDAYS**, WRITE "20" IF **ALL DAYS including weekends**, WRITE "30" IF ONE TIME PER WEEK, WRITE "4"	A. # OF DAYS PER MONTH IN FACILITY ____ B. # DAYS PER MONTH IN OUTREACH LOCATIONS ____	
402	Now I would like to ask you about the health personnel that work in this facility. I will read the type of health worker and for each one I would like you to tell me the number sanctioned by the Ministry of Health to work in this facility and the ones who are here today.		
	JOB OF HEALTH WORKER	A. # WORKERS **SANCTIONED** TO WORK IN THIS FACILITY (FULL OR PART-TIME) / B. # WORKERS WHO ARE **PRESENT TODAY**	
01	DOCTOR		
02	CLINICAL OFFICER		
03	REGISTERED NURSE		
04	REGISTERED MIDWIFE		
05	ENROLLED NURSE		
06	ENROLLED MIDWIFE		
07	LABORATORY TECHNICIAN		
08	LABORATORY ASSISTANT		
09	PHARMACIST		
10	OTHER CLINICAL CARE STAFF (NURSING ASSIATNTS, ETC.)		
11	ALL OTHER ASSIGNED STAFF (for instance, clerical staff, cleaning staff, etc.)		
403	During the past three years have you received any pre-service or in-service training on subjects related to maternal, child, or newborn health or illness?	YES ..... 1 NO 2	→405
404	Did you receive the training in any topic related to the following topics that I will read? IF YES, THEN ASK: When was your most recent training? READ THE LIST	YES, IN PAST 12 MONTHS / YES, IN PAST 2-3 YEARS / NO TRAINING WITHIN PAST 3 YEARS	
01	Management of fever in children	1 / 2 / 3	
01b	IF YES, ASK: Did this training cover the following topics		
	a) How to evaluate patients with fever	1 / 2 / 3	
	b) How to select patients for RDT testing	1 / 2 / 3	
	c) Performing and reading an RDT	1 / 2 / 3	
	d) Management of a patient with fever and a positive RDT	1 / 2 / 3	
	e) Management of a patient with fever and a negative RDT	1 / 2 / 3	
	f) Recognition and referral of patients with severe illness	1 / 2 / 3	
	g) Patient education	1 / 2 / 3	
	h) RDT storage and monitoring	1 / 2 / 3	
	i) Treatment with ACT	1 / 2 / 3	
01c	Did the training also cover the following topics?		
	a) Treatment of pneumonia or Acute Respiratory Infections	1 / 2 / 3	
	b) Diarrhea treatment	1 / 2 / 3	
	c) Diarrhea treatment	1 / 2 / 3	
	d) Integrated Management of Childhood Illness (IMCI)	1 / 2 / 3	
	e) Nutrition/breastfeeding (for instance, complementary feeding, micronutrients) (for instance, complementary feeding, micronutrients)	1 / 2 / 3	
	f) Breastfeeding	1 / 2 / 3	

405	Now I would like to ask you some questions about supervision from a supervisor outside the facility. a. Do you receive technical support or supervision in your work? b. IF YES, ASK: When was the most recent time?	YES, IN THE PAST 3 MONTHS YES, IN THE PAST 4-6 MONTHS YES, IN THE PAST 7-12 MONTHS YES, MORE THAN 12 MONTHS AGO NO SUPERVISION			1 2 3 4 5	→ →407 →407 →407	

406	The last time you were personally supervised, did your supervisor do any of the following? READ THE LIST:		YES	NO	DON'T KNOW	
01	Deliver supplies	DELIVERED SUPPLIES	1	2	9	
02	Check your records or reports	CHECKED RECORD	1	2	9	
03	Observe your work	OBSERVED	1	2	9	
04	Provide any feedback (either positive or negative) on your performance	GAVE FEEDBACK	1	2	9	
05	Provide any comment that you were **doing your work well**	GAVE PRAISE	1	2	9	
06	Provide updates on administrative or technical issues related to your work	GAVE UPDATES	1	2	9	
07	Discuss problems you have encountered	DISCUSSED PROBLEMS	1	2	9	
08	Checked drug supply	CHECKED DRUG SUPPLY	1	2	9	

407	When was the last time that this health facility referred a child with fever to a higher level facility?	IN LAST MONTH 1 - 3 MONTHS AGO 4 - 12 MONTHS AGO OVER A YEAR AGO NEVER	1 2 3 4 5	

ASK THE HEALTH WORKER TO IDENTIFY THE OUTPATIENT PATIENT CONSULTATION REGISTER FOR THE HEALTH FACILITY. DO NOT INCLUDE INPATIENT RECORDS. USE THE REGISTER (AND OTHER HEALTH RECORDS SUCH AS PATIENT CARD) TO ANSWER THE QUESTIONS BELOW.

408	Is there a sick child consultation register? IF YES, ASK TO SEE THE REGISTER	OBSERVED REGISTER REPORTED, NOT SEEN NO REGISTER	1 2 3	 →412 →412

REVIEW ENTRIES IN THE SICK CHILD REGISTER (ONLY THE ENTRIES FOR CHILDREN U5 IF ADULT AND U5 REGISTERS COMBINED). IDENTIFY THE ***LAST 20 NEW CASES*** OF <5YR. YEAR SICK CHILDREN SEEN AT THE HEALTH FACILITY WITH ANY OF THE FOLLOWING DIAGNOSES: FEVER/MALARIA, PNEUMONIA/FAST OR DIFFICULT BREATHING, DIARRHEA (without blood), OR SEVERE DISEASE. USE THE INFORMATION AVAILABLE ABOUT THESE CASES TO ANSWER THE QUESTIONS BELOW.

409	DOES THE REGISTER (AND OTHER HEALTH RECORDS, IF NEEDED) CONTAIN COMPLETE INFORMATION ON AGE, DIAGNOSIS, & TREATMENT ***FOR THE LAST 20 NEW CASES*** OF <5YR. SICK CHILDREN WITH EITHER MALARIA, PNEUMONIA, OR DIARRHEA?	01 NO. OF CASES WITH AGE DOCUMENTED 02 NO. OF CASES WITH DIAGNOSIS/CLASSIFICATION DOCUMENTED 03 NO. OF CASES WITH TREATMENT DOCUMENTED	
	RECORD THE NUMBER OF TIMES OUT OF 20 CASES THAT THE INFORMATION WAS DOCUMENTED FOR EACH CASE	04 NO. OF CASES WITH **ALL** OF ABOVE DOCUMENTED 05 NO. OF CASES WITH **NONE** OF ABOVE DOCUMENTED	
410	HOW RECENT IS THE DATE OF THE MOST RECENT ENTRY?	WITHIN THE PAST 7 DAYS MORE THAN 7 DAYS OLD	1 2

411	BREAKDOWN OF **LAST 20 NEW SICK CHILD CASES**.....					
01	MALARIA OR FEVER (UNCOMPLICATED)	A1. NO. OF NEW MALARIA CASES OF U5 CHILDREN AMONG THE 20		A2	NO. MALARIA CASES TREATED WITH ACT	
02	PNEUMONIA OR ACUTE RESPIRATORY INFECTION (ARI)	B1 NO. OF PNEUMONIA/ARI CASES OF U5 CHILDREN AMONG THE 20		B2	NO. PNEUMONIA CASES TREATED WITH COTRIMOXAZOLE	
03	DIARRHEA WITHOUT BLOOD	C1. NO. OF DIARRHEA CASES OF U5 CHILDREN AMONG THE 20		C2	NO. DIARRHEA CASES TREATED WITH ORS & **NO ANTIBIOTIC**	
04	SEVERE DISEASE *(SEVERE MALARIA, SEVERE PNEUMONIA, SEVERE FEBRILE ILLNESS)*	D1. NO. OF SEVERE DISEASE CASES OF CHILDREN U5 AMONG THE 20		D2	NO. SEVERE DISEASE CASES REFERRED TO NEXT HIGHER LEVEL FACILITY	

412	Can you please show me a copy of the latest monthly service report that you sent to the District Health Office? EXAMINE THE REPORT	LATEST REPORT SEEN AND **LESS THAN 3 MONTHS** OLD LATEST REPORT SEEN AND **OLDER THAN 3 MONTHS** OLD REPORT SAID TO BE LESS THAN 3 MONTHS, NOT OBSERVED REPORT SAID TO BE MORE THAN 3 MONTHS, NOT OBSERVED NO REPORT	1 2 3 4 5	

413	LOOK FOR EVIDENCE OF USE OF SERVICE DATA Can you tell me if you have a wall chart or graphs or have had a meeting among the health facility staff to discuss the monthly service report **(MSR)** data within the last 3 months? CIRCLE **ALL** THAT APPLY	WALL CHART SUMMARIZING MSR DATA GRAPH SUMMARIZING MSR DATA MEETING TO DISCUSS MSR DATA IN IN LAST 3 MO. OTHER: SPECIFY NONE OF THE ABOVE	A B C D E	

414 For each of the following diagnostic tests, please tell me if this facility can conduct the test and has all items so it can be done today, or if the facility has a system for having the test conducted elsewhere but getting results returned for follow up by this facility.

	DIAGNOSTIC TEST	YES, CAN BE CONDUCTED ONSITE TODAY	YES, OBSERVED SYSTEM FOR TEST OUTSIDE	TEST NOT AVAIL-ABLE
01	Complete blood count	1	2	3
02	Anemia (hemoglobinhematocrit, or litmus paper)	1	2	3
03	Malaria (rapid test or microscopy)	1	2	3
04	Urine glucose (dipstix or benedicts test)	1	2	3
05	Urine protein (dipstix or acetic acid)	1	2	3
06	HIV (rapid, ELISA, or Western Blott)	1	2	3
07	AFB for TB	1	2	3
08	Syphilis (VDRL or RPR)	1	2	3

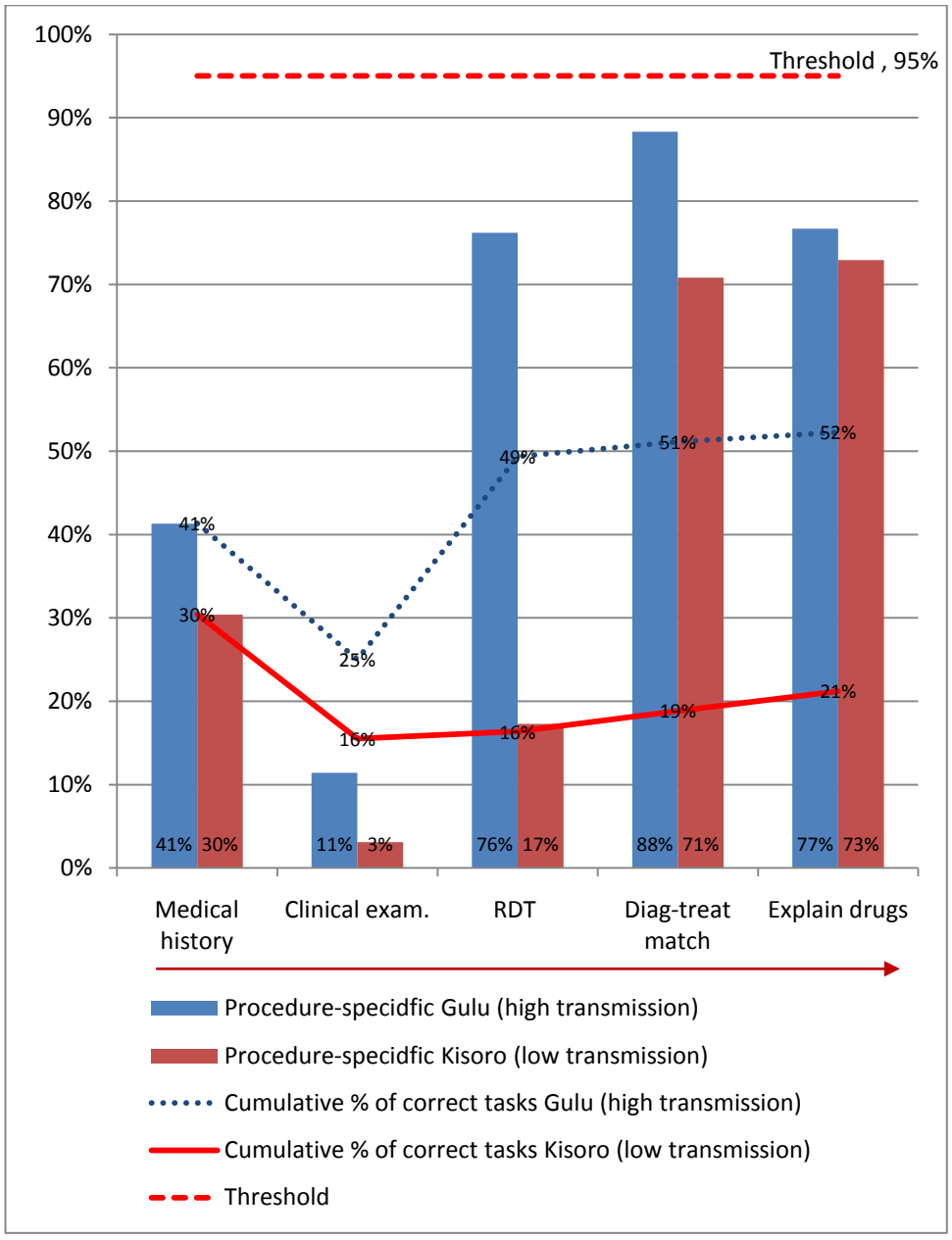
NOTE ANY QUALITATIVE OBSERVATIONS HERE:

Appendix 19 Adherence scores for specific tasks by facilities

		MEDICAL HISTORY						CLINICAL EXAMINATION						PERFORMING RAPID DIAGNOSTIC TEST											TREATMENT	COUNSELLING
		Success rate for History taking (DR = 6:5)						Success rate for Physical examination (DR = 6:5)						Success rate for using RDT (DR = 6:5)											DR = 6:5	DR = 6:5
District	LQAS No.	HW_ask_fever	HW_ask_cough or difficulty in	Diarrhoea	HW_ask_feeding	HW_ask_vomiting	HW_ask_convulsions	Vaccination status	Dehydration	Respiration rate	Conjuctiva for palor and jaundice	Throat for redness	Ear for discharge or sores	Requested RDT	Checks expeiry date	Puts on gloves	Writes Patents' names on casette	Cleans patients' fingers	Allows finger to dry before pricking	Uses loop or pipette to collect blood	Puts blod into position A	Puts buffer in position B	Waits at at least 15 mins after buffer	Communicates RDT results	Exp dia-tx match per facility	Expected no of correct tasks per facility
GULU	A1	P	P	F	F	F	F	F	F	F	F	F	F	F	F	F	F	F	F	F	F	F	F	F	P	P
(n = 10)	A2	P	P	F	F	F	F	F	F	F	F	F	F	P	F	P	F	P	F	P	P	P	F	P	P	P
	A3	P	P	F	P	F	F	F	F	F	F	F	F	P	P	P	F	P	P	P	P	P	P	P	F	F
	B1	P	F	F	F	F	F	F	F	F	F	F	F	P	P	F	P	P	F	P	P	P	P	F	P	F
	B2	P	P	F	F	F	F	F	F	F	F	F	F	P	F	P	F	P	F	P	P	P	F	P	P	P
	B3	P	P	F	F	F	F	F	F	F	F	F	F	P	P	P	P	P	P	P	P	P	P	P	P	F
	C1	P	F	F	F	F	F	F	F	F	F	F	F	P	F	P	P	P	F	P	P	P	F	P	P	P
	C2	P	P	P	F	F	F	F	F	F	F	F	F	P	P	P	F	P	F	P	P	P	F	P	F	P
	C3	P	P	F	F	P	F	F	F	F	F	F	F	P	P	P	P	P	P	P	P	P	P	P	P	F
	D	P	P	P	F	F	F	F	F	F	F	F	F	P	P	P	P	P	F	P	P	P	F	P	F	P
Classification	Number of successes	10	8	2	1	1	0	0	0	0	0	0	0	9	6	8	5	9	3	9	9	9	4	8	7	6
of district																										
	Deceision rule: 7	H	H	L	L	L	L	L	L	L	L	L	L	H	L	H	L	H	L	H	H	H	L	H	H	L
KISORO	KA	P	F	P	F	F	F	F	F	F	F	F	F	F	F	F	F	F	F	F	F	F	F	F	F	P
(n = 8)	KB	P	P	P	F	F	F	F	F	F	F	F	F	F	F	F	F	F	F	F	F	F	F	F	P	P
	KC	P	F	F	F	F	F	F	F	F	F	F	F	P	F	P	P	P	F	P	P	P	F	P	F	P
	KE	P	F	F	F	F	F	F	F	F	F	F	F	F	F	F	F	F	F	F	F	F	F	F	P	P
	KF	P	F	F	F	F	F	F	F	F	F	F	F	F	F	F	F	F	F	F	F	F	F	F	P	F
	KG	P	P	F	F	F	F	F	F	F	F	F	F	F	F	F	F	F	F	F	F	F	F	F	F	P
	KH	P	P	P	F	F	F	F	F	F	F	F	F	F	F	F	F	F	F	F	F	F	F	F	F	F
	KI	P	P	F	F	F	F	F	F	F	F	F	F	F	F	F	F	F	F	F	F	F	F	F	P	P
Classification	Number of successes	8	4	3	0	0	0	0	0	0	0	0	0	1	0	1	1	1	0	1	1	1	0	1	4	6
of district																										
	Decision rule: 6	H	L	L	L	L	L	L	L	L	L	L	F	L	L	L	L	L	L	L	L	L	L	F	L	H

LQAS Numbers (A1 to D in Gulu, and KA to KI in Kisoro) are codes assigned to specific HCs that were enrolled into the survey

Appendix 20 Procedure-specific and cumulative adherence scores by successive steps



Bibliography

1. World Health Organisation, *World Malaria Report*, in *WHO/HTM/GMP/2008.1*2008.

2. World Health Organisation and Unicef, in *TheAfrica Malaria Report 2003*2003.

3. Uganda Ministry of Health, *Uganda Malaria Control Strategic Plan 2005/06 - 2009/10*, M.C.P. Ministry of Health (Uganda), Editor undated. p. 8.

4. World Bank/UN Common database, *Globalis, Uganda*.

5. Gallup, J. and J. Sachs, *The economic burden of malaria.* American Journal of Tropical Medicine and Hygiene, 2001. **64**(1_suppl): p. 85-96.

6. World Health Organisation Roll Back Malaria, *Country Profile: Uganda.* World Malaria Report, 2005, 2005.

7. Presidential Malaria Initiative (PMI)-Uganda, *Malaria Operational Plan (MOP) FY 2007* 2007.

8. Lin, D. and H. Tai, *To Donate Now or Later? Estimating the Time Value of Social Impact* The Social Impact Consulting Think Tank Journal 2008. **Spring 2008 papers**.

9. World Health Organisation, *Guidelines for the Treatment of Malaria*, 2006.

10. World Health Organisation, *Malaria epidemics: forecasting, prevention early detection and control - from policy to practice*, G. World Health Organisation, Editor 2004.

11. Koram, K.A. and M.E. Molyneux, *When Is "Malaria" Malaria? The Different Burdens of Malaria Infection, Malaria Disease, and Malaria-Like Illnesses.* The American journal of tropical medicine and hygiene, 2007. **77**(6 Suppl): p. 1-5.

12. Rolland, E., et al., *Operational response to malaria epidemics: Are rapid diagnostic tests cost-effective?* Tropical Medicine and International Health, 2006. **11**(4): p. 398-408.

13. Zikusooka, C.M., D. McIntyre, and K.I. Barnes, *Should countries implementing an artemisinin-based combination malaria treatment policy also introduce rapid diagnostic tests?* Malaria Journal, 2008. **7**.

14. Bell, D. *Is there a role for Rapid Diagnostic Tests in Africa?* . 2004.

15. Malenga, G., et al., *Antimalarial treatment with artemisinin combination therapy in Africa.* British Medical Journal 2005. **331**(7519): p. 706-7.

16. World Health Organisation, *WHO briefing on Malaria Treatment Guidelines and artemisinin monotherapies Geneva, 19*, W.H. Organisation, Editor 2006: Geneva.

17. Taylor, W.R.J., et al., *Use of weight-for-age-data to optimize tablet strength and dosing regimens for a new fixed-dose artesunate-amodiaquine combination for treating falciparum malaria.* Bulletin of the World Health Organization, 2006. **84**: p. 956-964.

18. Obua, C., *Fixed-dose Chloroquine and Sulphadoxine/Pyrimethamine Treatment of Malaria: Outcome and Pharkokinetics Aspects*, 2007, Makaere University and University of Karolinska.

19. World Health Organisation, *"Guidelines for the Treatment of Malaria." WHO Library Cataloguing-in-Publication Data.*, 2010.

20. Wongsrichanalai, C., et al., *A review of malaria diagnostic tools: microscopy and rapid diagnostic test (RDT).* American Journal of Tropical Medicine & Hygiene, 2007. **77**(6 Suppl): p. 119-27.

21. Murray, C.K. and J.W. Bennett, *Rapid Diagnosis of Malaria.* Interdisciplinary Perspectives on Infectious Diseases, 2009. **2009**: p. 7.

22. Murray, C.K., et al., *Update on Rapid Diagnostic Testing for Malaria.* Clinical Microbiology Reviews, 2008. **21**(1): p. 97-110.

23. Guerin, P.J., et al., *Malaria: current status of control, diagnosis, treatment, and a proposed agenda for research and development.* The Lancet Infectious Diseases, 2002. **2**(9): p. 564-73.

24. Moody, A., *Rapid diagnostic tests for malaria parasites.* Clinical Microbiology Reviews, 2002. **15**(1): p. 66-78.

25. Olliaro, P., *Management of fever and malaria - policy and practice.* Trop Med Int Health, 2009.

26. English, M., et al., *Abandoning Presumptive Antimalarial Treatment for Febrile Children Aged Less Than Five Yearsâ€"A Case of Running Before We Can Walk?* PLoS Med, 2009. **6**(1): p. e1000015.

27. Campbell, M., et al., *Framework for design and evaluation of complex interventions to improve health.* British Medical Journal 2000. **321**(7262): p. 694-6.

28. Rutten, G., et al., *A theory-based cross-sectional survey demonstrated the important role of awareness in guideline implementation.* Journal of Clinical Epidemiology, 2009. **62**(2): p. 167-176.e1.

29. Wensing, M., T. van der Weijden, and R. Grol, *Implementing guidelines and innovations in general practice: which interventions are effective?* British Journal of General Practice, 1998. **48**(427): p. 991-7.

30. Grol, R., et al., *Attributes of clinical guidelines that influence use of guidelines in general practice: observational study.* British Medical Journal 1998. **317**(7162): p. 858-861.

31. Harting, J., et al., *A qualitative application of the diffusion of innovations theory to examine determinants of guideline adherence among physical therapists.* PhysIcal Therapy, 2009. **89**(3): p. 221-32.

32. Reyburn, H., et al., *Rapid diagnostic tests compared with malaria microscopy for guiding outpatient treatment of febrile illness in Tanzania: randomised trial.* British Medical Journal 2007. **334**(7590): p. 403.

33. Zurovac, D., et al., *Modeling the financial and clinical implications of malaria rapid diagnostic tests in the case-management of older children and adults in Kenya.* The American Journal of Tropical Medicine & Hygiene, 2008. **78**(6): p. 884-91.

34. Bisoffi, Z., et al., *Rapid malaria diagnostic tests vs. clinical management of malaria in rural Burkina Faso: safety and effect on clinical decisions. A randomized trial.* Tropical Medicine & International Health, 2009. **14**(5): p. 491-498.

35. Skarbinski, J., et al., *Effect of malaria rapid diagnostic tests on the management of uncomplicated malaria with artemether-lumefantrine in Kenya: a cluster randomized trial.* American Journal of Tropical Medicine & Hygiene, 2009. **80**(6): p. 919-26.

36. Ansah, E.K., et al., *Rapid testing for malaria in settings where microscopy is available and peripheral clinics where only presumptive treatment is available: a randomised controlled trial in Ghana.* British Medical Journal 2010. **340**: p. c930.

37. Yeboah-Antwi, K., et al., *Community case management of fever due to malaria and pneumonia in children under five in Zambia: a cluster randomized controlled trial.* PLoS Med, 2010. **7**(9): p. e1000340.

38. Richard Grol, M.W., *What drives change? Barriers to and incentives for achieving evidence-based practice.* Medical Journal of Australia, 2004. **180**(6 Suppl): p. S57.

39. Eccles, M., et al., *Changing the behavior of healthcare professionals: the use of theory in promoting the uptake of research findings.* Journal of Clinical Epidemiology, 2005. **58**(2): p. 107-112.

40. Flay, B.R., *Efficacy and effectiveness trials (and other phases of research) in the development of health promotion programs.* Preventive Medicine, 1986. **15**(5): p. 451-74.

41. Orlandi, M.A., *The diffusion and adoption of worksite health promotion innovations: An analysis of barriers.* Preventive Medicine, 1986. **15**(5): p. 522-536.

42. Rogers E, *Diffusion of Innovations*. 1st ed1962, New York, London: The Free Press of Glencoe; Macmillan New York.

43. Grol, R. and J. Grimshaw, *From best evidence to best practice: effective implementation of change in patients' care.* The Lancet, 2003. **362**(9391): p. 1225-1230.

44. Rycroft-Malone, J., et al., *An exploration of the factors that influence the implementation of evidence into practice.* Journal of Clinical Nursing, 2004. **13**(8): p. 913-24.

45. Graham, I.D., et al., *Lost in knowledge translation: Time for a map?* Journal of Continuing Education in the Health Professions, 2006. **26**(1): p. 13-24.

46. Greenhalg, T., et al., *Diffusion of Innovations in Service Organizations: Systematic Review and Recommendations.* The Milbank Quarterly, 2004. **82**(4): p. 581-629.

47. Greer, A., *Medical technology: Assessment, adoption, and utilization.* Journal of Medical Systems, 1981. **5**(1): p. 129-145.

48. Ward, V., A. House, and S. Hamer, *Developing a framework for transferring knowledge into action: a thematic analysis of the literature.* Journal of Health Services Research & Policy, 2009. **14**(3): p. 156-164.

49. Prochaska, J.O., et al., *Efficacy and Effectiveness Trials: examples from smoking cessation and bullying prevention.* Journal of Health Psychology, 2007. **12**(1): p. 170-178.

50. Streiner, D.L., *The 2 "Es" of research: efficacy and effectiveness trials.* Canadian Journal Psychiatry, 2002. **47**(6): p. 552-6.

51. Gafni, A. and S. Birch, *Guidelines for the adoption of new technologies: a prescription for uncontrolled growth in expenditures and how to avoid the problem.* Canadian Medical Assiociation Journal, 1993. **148**(6): p. 5.

52. Laupacis, A., et al., *How attractive does a new technology have to be to warrant adoption and utilization? Tentative guidelines for using clinical and economic evaluations.* Canadian Medical Assiociation Journal, 1992. **146**(4): p. 473-81.

53. World Health Organisation, *The role of laboratory diagnosis to support malaria disease management; focus on the use of rapid diagnostic tests in areas of high transmission* in *Report of the WHO technical consultation, 25 - 26 October 2004*2006.

54. Grol, R. and M. Wensing, *What drives change? Barriers to and incentives for achieving evidence-based practice.* Medical Journal of Australia, 2004. **180**(6 Suppl): p. S57-60.

55. Berwick, D.M., *Disseminating Innovations in Health Care.* The Journal of American Medical Association, 2003. **289**(15): p. 1969-1975.

56. Ramanarayanan, S. *Does Practice Make Perfect: An Empirical Analysis of Learning-by-Doing in Cardiac Surgery (April 2008).* Working paper series, UCLA Anderson School of Management, 2008.

57. Wiersma, E., *Conditions That Shape the Learning Curve: Factors That Increase the Ability and Opportunity to Learn.* MANAGEMENT SCIENCE, 2007. **53**(12): p. 1903-1915.

58. Benkard, C.L., *Learning and forgetting: the dynamics of aircraft production.* American Economics Review, 2000. **90**(4): p. 20.

59. Boone, G. and Hicks, *Learning and Knowledge Depreciation in Professional Services.* MANAGEMENT SCIENCE, 2000. **54**(7): p. 6.

60. Shafer, S.M., D.A. Nembhard, and M.V. Uzumeri, *The Effects of Worker Learning, Forgetting, and Heterogeneity on Assembly Line Productivity.* MANAGEMENT SCIENCE, 2001. **47**(12): p. 1639-1653.

61. Huesch, M.D. and M. Sakakibara, *Forgetting the learning curve for a moment: how much performance is unrelated to own experience?* Health Economics, 2009. **18**(7): p. 855-862.

62. Rowe, S.Y., et al., *Longitudinal analysis of community health workers' adherence to treatment guidelines, Siaya, Kenya, 1997-2002.* Trop Med Int Health, 2007. **12**(5): p. 651-63.

63. Britten, N. and O. Ukoumunne, *The influence of patients' hopes of receiving a prescription on doctors' perceptions and the decision to prescribe: a questionnaire survey.* British Medical Journal 1997. **315**(7121): p. 1506-1510.

64. Chandler, C.I.R., et al., *Malaria overdiagnosis: is patient pressure the problem?* Health Policy Plan., 2008: p. czm046.

65. World Health Organisation, et al., *Malaria rapid diagnostic test performance. Results of WHO product testing of malaria RDTs: rounds 3 (2010-2011)*, 2011, World Health Organisation, Special programme for research and training in tropical diseases.

66. Shiff, C.J., Z. Premji, and J.N. Minjas, *The rapid manual ParaSight-F test. A new diagnostic tool for Plasmodium falciparum infection.* Trans R Soc Trop Med Hyg, 1993. **87**(6): p. 646-8.

67. Mason, D.P., et al., *A comparison of two rapid field immunochromatographic tests to expert microscopy in the diagnosis of malaria.* Acta Trop, 2002. **82**(1): p. 51-9.

68. Bell, D. and R.W. Peeling, *Evaluation of rapid diagnostic tests: malaria.* Nature Reviews Microbiology.

69. World Health organisation, W.P.r., *Towards quality testing of malaria rapid diagnostic tests: evidence and methods.*, 2006, World Health organisation, Western Pacific region, Manilla, Philippines.

70. D'Acremont, V., et al., *Time to move from presumptive malaria treatment to laboratory-confirmed diagnosis and treatment in African children with fever.* PLoS Med, 2009. **6**(1): p. e252.

71. Talisuna, A.O. and D.N. Meya, *Diagnosis and treatment of malaria.* British Medical Journal 2007. **334**(7590): p. 375-6.

72. Shillcutt, S., 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): p. 101-110.

73. Tucker, J., et al. *Health demonstration projects: Evaluating a community-based health intervention programme to improve young people's sexual health*. Critical Public Health, 2006. **16**, 175-189 DOI: 10.1080/09581590600986358.

74. Linnan, L., Steckler, A.,, *An overview*, in *Process evaluation for public health interventions and research*, A. Steckler and L. Linnan, Editors. 2002, John Wiley & Sons. p. 432.

75. Okello, P.E., et al., *Variation in malaria transmission intensity in seven sites throughout Uganda.* American Journal of Tropical Medicine and Hygiene, 2006. **75**(2): p. 219-225.

76. Beier, J.C., G.F. Killeen, and J.I. Githure, *Short report: entomologic inoculation rates and Plasmodium falciparum malaria prevalence in Africa.* The American journal of tropical medicine and hygiene, 1999. **61**(1): p. 109-13.

77. Mendis, K., et al., *From malaria control to eradication: The WHO perspective.* Trop Med Int Health, 2009. **14**(7): p. 802-9.

78. Roll Back Malaria *Epidemiology of malaria*. undated.

79. Hay, S.I., D.L. Smith, and R.W. Snow, *Measuring malaria endemicity from intense to interrupted transmission.* Lancet Infect Dis, 2008. **8**(6): p. 369-78.

80. Hopkins, H., et al., *Rapid Diagnostic Tests for Malaria at Sites of Varying Transmission Intensity in Uganda.* The Journal of Infectious Diseases, 2008. **197**(4): p. 510-518.

81. Dorsey, G. and H. Hopkins, *Effectiveness and safety of training in fever case management incorporating rapid diagnostic tests (RDTs) for malaria at peripheral health centers in Uganda; draft version 1.0*, 2007, Uganda Malaria Surveillance Project (UMSP).

82. World Health Organisation, *Quality of care: a process of making strategic choices in health systems*. Vol. W. 84.1. 2006, Geneva, Switzerland.

83. Coyle, Y.M. and J.M. Battles, *Using antecedents of medical care to develop valid quality of care measures.* International Journal for Quality in Health Care, 1999. **11**(1): p. 5-12.

84. Donebanian, A., *Explorations in Quality Assessment and Monitoring. The Definition of Quality and Approaches to its Assessment*. Vol. 1. 1980: Ann Abbor, MI: Health Administration Press.

85. Brook, R.H., E.A. McGlynn, and P.D. Cleary, *Quality of health care. Part 2: measuring quality of care.* The New England Journal of Medicine, 1996. **335**(13): p. 966-70.

86. The National Roundtable on Health Care, Q. and M. Institute of, *Measuring the Quality of Health Care*, ed. M.S. Donaldson1999: The National Academies Press.

87. Edlund, M.J., et al., *Does satisfaction reflect the technical quality of mental health care?* Health Serv Res, 2003. **38**(2): p. 631-45.

88. Tabrizi, J.S., et al., *Patient perspectives on consistency of medical care with recommended care in type 2 diabetes.* Diabetes Care, 2007. **30**(11): p. 2855-6.

89. Kenagy Jw, B.D.M.S.M.F., *Service quality in health care.* JAMA: The Journal of the American Medical Association, 1999. **281**(7): p. 661-665.

90. Naimoli, J.F., et al., *Effect of the Integrated Management of Childhood Illness strategy on health care quality in Morocco.* International Journal for Quality in Healthcare, 2006. **18**(2): p. 134-44.

91. Osterholt, D., et al., *Predictors of treatment error for children with uncomplicated malaria seen as outpatients in Blantyre district, Malawi.* Tropical Medicine & International Health, 2006. **11**(8): p. 1147 - 1156.

92. Pariyo, G.W., et al., *Improving facility-based care for sick children in Uganda: training is not enough.* Health Policy Plan., 2005. **20**(suppl_1): p. i58-68.

93. Rowe, S.Y., et al., *Effect of multiple interventions on community health workers' adherence to clinical guidelines in Siaya district, Kenya.* Trans R Soc Trop Med Hyg, 2007. **101**(2): p. 188-202.

94. Gross, R., et al., *The relationship between primary care physicians' adherence to guidelines for the treatment of diabetes and patient satisfaction: findings from a pilot study.* Family Practice, 2003. **20**(5): p. 563-569.

95. Valk, G.D., et al., *Quality of Care for Patients with Type 2 Diabetes Mellitus in the Netherlands and the United States: A Comparison of Two Quality Improvement Programs.* Health Services Research, 2004. **39**(4p1): p. 709-726.

96. Clancy, D.E., et al., *Evaluating concordance to American Diabetes Association standards of care for type 2 diabetes through group visits in an uninsured or inadequately insured patient population.* Diabetes Care, 2003. **26**(7): p. 2032-6.

97. Sibbald, B. and M. Roland, *Understanding controlled trials: Why are randomised controlled trials important?* BMJ, 1998. **316**.

98. Sanson-Fisher, R.W., et al., *Limitations of the Randomized Controlled Trial in Evaluating Population-Based Health Interventions.* American Journal of Preventive Medicine, 2007. **33**(2): p. 155-161.

99. Higgins, J.P.T. and S. Green, *Cochrane Handbook for Systematic Reviews of Interventions; Version 5.1.0 (updated March 2011); The Cochrane Collaboration*. Version 5.1.0 ed, ed. J.P.T. Higgins and S. Green2011: John Wiley and Sons, Ltd, 2011.

100. Stolberg, H.O., G. Norman, and I. Trop, *Randomized Controlled Trials.* American Journal of Roentgenology, 2004. **183**(6): p. 1539-1544.

101. Rosner A., B., *Fundamentals of biostatistics.* 5th ed2000, Duxbury, England: Thompson Learning.

102. Rowe, A.K., et al., *Design effects and intraclass correlation coefficients from a health facility cluster survey in Benin.* International Journal for Quality in Healthcare, 2002. **14**(6): p. 521-3.

103. Sally, M.K. and J.M. Bland, *Statistics notes: Sample size in cluster randomisation.* BMJ, 1998. **316**.

104. Valadez, J.J., *Assessing Child Survival Programs in Developing Countries: Testing Lot Quality Assurance Sampling.* Havard Series on Population and International Health1991: Cambridge, M.A. Havard University Press.

105. Valadez, J.J. and R.D. Babu, eds. *Decentralised Supervision of Community Health Programmes: Usinn LQAS in Two Districts of South Nepal.* Community-Based Health Care: Lessons from Bangladesh to Boston, ed. J.R.a.J. Wyon2002, Management Sciences for Health Publications, in collaboration with the Harvard School of Public Health: Boston.

106. Valadez, J.J., et al., *Assessing family planning service-delivery skills in Kenya.* Studies In Family Planning, 1997. **28**(2): p. 143-50.

107. Robertson, S.E. and J.J. Valadez, *Global review of health care surveys using lot quality assurance sampling (LQAS), 1984-2004.* Social Science & Medicine, 2006. **63**(6): p. 1648-1660.

108. Rowe, A.K., et al., *Predictors of Correct Treatment of Children with Fever Seen at Outpatient Health Facillities in the Central African Republic.* American Journal of Epidemiology, 2000. **151**(10): p. 1029-1035.

109. Boonstra, E., et al., *Adherence to treatment guidelines in primary health care facilities in Botswana.* Trop Med Int Health, 2002. **7**(2): p. 178-86.

110. Trap, B., et al., *The impact of supervision on stock management and adherence to treatment guidelines: a randomized controlled trial.* Health Policy Plan, 2001. **16**(3): p. 273-80.

111. Zurovac, D., et al., *Effects of revised diagnostic recommendations on malaria treatment practices across age groups in Kenya.* Trop Med Int Health, 2008. **13**(6): p. 784-7.

112. Zurovac, D., et al., *Predictors of the quality of health worker treatment practices for uncomplicated malaria at government health facilities in Kenya.* International Journal Of Epidemiology, 2004. **33**(5): p. 1080-1091.

113. Krause, G., et al., *From diagnosis to drug taking: staff compliance with guidelines and patient compliance to prescriptions in Burkina Faso.* International Journal for Quality in Healthcare, 2000. **12**(1): p. 25-30.

114. Krause, G., et al., *Diagnostic quality in rural health centres in Burkina Faso.* Tropical Medicine & International Health, 1998. **3**(2): p. 100-107.

115. Rowe, A.K., et al., *Management of Childhood Illness at Health Facilities in Benin: Problems and Their Causes.* American Journal of Public Health, 2001. **91**(10): p. 1625-1635.

116. Rowe, A.K., et al., *Risk and protective factors for two types of error in the treatment of children with fever at outpatient health facilities in Benin.* International Journal Of Epidemiology, 2003. **32**(2): p. 296-303.

117. Rowe, A.K., et al., *A multifaceted intervention to improve health worker adherence to integrated management of childhood illness guidelines in Benin.* American Journal of Public Health, 2009. **99**(5): p. 837-46.

118. Baiden, F., et al., *An Evaluation of the Clinical Assessments of Under-Five Febrile Children Presenting to Primary Health Facilities in Rural Ghana.* PLoS ONE, 2011. **6**(12): p. e28944.

119. Higgins, J.P., et al., *The Cochrane Collaboration's tool for assessing risk of bias in randomised trials.* British Medical Journal 2011. **343**: p. d5928.

120. Kyabayinze, D.J., et al., *Use of RDTs to improve malaria diagnosis and fever case management at primary health care facilities in Uganda.* Malaria Journal, 2010. **9**: p. 200.

121. Msellem, M.I., et al., *Influence of Rapid Malaria Diagnostic Tests on Treatment and Health Outcome in Fever Patients, Zanzibarâ€"A Crossover Validation Study.* PLoS Med, 2009. **6**(4): p. e1000070.

122. Bonate, L.P., *Aanalysis of pretest-posttest designs*2000: Chapman and Hall/CRC.

123. Dimitrov, D.M. and P.D. Rumrill, Jr., *Pretest-posttest designs and measurement of change.* Work, 2003. **20**(2): p. 159-65.

124. World Health Organisation, *Technical updates of the guidelines on the Integrated Management of Childhood Illness (IMCI): evidence and recommendations for further adaptations*, 2005.

125. Garner, P. and G.D. Smith, *Information for decision making: lot quality assurance sampling in the spotlight.* International Journal Of Epidemiology, 2010. **39**(1): p. 5-6.

126. Brooker, S., et al., *Rapid assessment of Schistosoma mansoni: the validity, applicability and cost-effectiveness of the Lot Quality Assurance Sampling method in Uganda.* Tropical Medicine & International Health, 2005. **10**(7): p. 647-658.

127. Valadez, J.J., et al., *Using lot quality assurance sampling to assess measurements for growth monitoring in a developing country's primary health care system.* International Journal of Epidemiology, 1996. **25**(2): p. 381-7.

128. Valadez, J.J., L. Weld, and W.V. Vargas, *Monitoring community health workers' performance through lot quality-assurance sampling.* American Journal of Public Health, 1995. **85**(8 Pt 1): p. 1165-6.

129. Lanata, C.F., et al., *An evaluation of lot quality assurance sampling to monitor and improve immunization coverage.* International Journal Of Epidemiology, 1990. **19**: p. 1086–90.

130. Rabarijaona, L., et al., *Validity of Lot Quality Assurance Sampling to optimize falciparum malaria surveys in low-transmission areas.* Transactions of the Royal Society of Tropical Medicine and Hygiene, 2001. **95**(3): p. 266-269.

131. Rabarijaona, L.P., et al., *Lot quality assurance sampling for screening communities hyperendemic for Schistosoma mansoni.* Tropical Medicine & International Health, 2003. **8**(4): p. 322-328.

132. Gerstl, S., et al., *Prevalence study of yaws in the Democratic Republic of Congo using the lot quality assurance sampling method.* PLoS ONE, 2009. **4**(7): p. e6338.

133. Biedron, C., et al., *An assessment of Lot Quality Assurance Sampling to evaluate malaria outcome indicators: extending malaria indicator surveys.* International Journal Of Epidemiology, 2010. **39**(1): p. 72-79.

134. Iboro, J., *A review of malaria diagnostic methods*, 2002, Liverpool Cchool of Tropical Medicine.

135. Ochola, L.B., et al., *The reliability of diagnostic techniques in the diagnosis and management of malaria in the absence of a gold standard.* Lancet Infect Dis, 2006. **6**(9): p. 582-8.

136. Bell, D. and R.W. Peeling, *Evaluation of rapid diagnostic tests: malaria.* Nat Rev Micro.

137. Murray, C.K., et al., *Rapid diagnostic testing for malaria.* Tropical Medicine & International Health, 2003. **8**(10): p. 876-883.

138. Graz, B., et al., *"Test and treat" or presumptive treatment for malaria in high transmission situations? A reflection on the latest WHO guidelines.* Malaria Journal, 2011. **10**: p. 136.

139. World Health Organisation, *New Perspectives, Malaria Diagnosis.* Report of a Joint WHO/USAID Informal Consultation 25–27 October 999, 2000. **WHO/CDS/RBM/2000.14; WHO/MAL/2000.1091.**

140. Lubell, Y., et al., *An interactive model for the assessment of the economic costs and benefits of different rapid diagnostic tests for malaria.* Malaria Journal, 2008. **7**: p. 21.

141. Goodman, C., P. Coleman, and A. Mills, *Economic analysis of malaria control in Sub-Saharan Africa*, G.f.f.h. research, Editor 2000, Global forum for health research: Geneva.

142. World Health Organisation, *World malaria report*, 2011.

143. Drummond, M., et al., *Methods for the economic evaluation of health care programmes. Third edition: Oxford University Press*2005.

144. Lubell, Y., et al., *The impact of response to the results of diagnostic tests for malaria: cost-benefit analysis.* British Medical Journal 2008. **336**(7637): p. 202-5.

145. Uzochukwu, B.S., et al., *Cost-effectiveness analysis of rapid diagnostic test, microscopy and syndromic approach in the diagnosis of malaria in Nigeria: implications for scaling-up deployment of ACT.* Malaria Journal, 2009. **8**: p. 265.

146. Rowe, A., et al., *Risk and protective factors for two types of error in the treatment of children with fever at outpatient health facilities in Benin.* International Journal of Epidemiology, 2003. **32**(2): p. 296 - 303.

147. Paoli, B., L.M. Haggard, and G. Shah *Confidence intervals in public health*. 2010. 1-12.

148. The Chochrane Collaboration, *RevMan 5 Userguide; Chocrane Information Management System* T.C. collarboration, Editor 2011, Chocrane Information Management System.

149. Osterholt, D.M., et al., *Improving pneumonia case-management in Benin: a randomized trial of a multi-faceted intervention to support health worker adherence to Integrated Management of Childhood Illness guidelines.* Hum Resour Health, 2009. **7**: p. 77.

150. Rowe, A., et al., *Predictors of correct treatment of children with fever seen at outpatient health facilities in the Central African Republic.* Am J Epidemiol, 2000. **151**(10): p. 1029 - 1035.

151. Chinkhumba, J., et al., *Comparative field performance and adherence to test results of four malaria rapid diagnostic tests among febrile patients more than five years of age in Blantyre, Malawi.* Malaria Journal, 2010. **9**(1): p. 209.

152. Faucher, J.F., et al., *Can treatment of malaria be restricted to parasitologically confirmed malaria? A school-based study in Benin in children with and without fever.* Malaria Journal, 2010. **9**.

153. Hillman, A.L., M.V. Pauly, and J.J. Kerstein, *How Do Financial Incentives Affect Physicians' Clinical Decisions and the Financial Performance of Health Maintenance Organizations?* New England Journal of Medicine, 1989. **321**(2): p. 86-92.

154. Boyd Cm, D.J.B.C.F.L.P.B.L.W.A.W., *Clinical practice guidelines and quality of care for older patients with multiple comorbid diseases: Implications for pay for performance.* JAMA: The Journal of the American Medical Association, 2005. **294**(6): p. 716-724.

155. Baiden, F., et al., *Would rational use of antibiotics be compromised in the era of test-based management of malaria?* Tropical Medicine & International Health, 2011. **16**(2): p. 142-144.

156. D'Acremont, V., et al., *Reduction of anti-malarial consumption after rapid diagnostic tests implementation in Dar es Salaam: a before-after and cluster randomized controlled study.* Malaria Journal, 2011. **10**: p. 107.

157. Uganda Ministry of Health Malaria Control Programme, *User's Manual: Use of rapid diagnostic tests (RDTs) for malaria in fever case management in Uganda; Near-final draft: 31 January 2009*, 2009.

158. Uganda Ministry of Health, *Uganda clinical guideline 2010; national guidelines on management of common conditions*, Q. assurance, Editor 2010, Ministry of Health, Uganda: Kampala.

159. National Medical Stores, *National Medical Stores catalogue and price indicator of 2008* N.M. Stores, Editor 2008.

160. Bank of Uganda, *Annual report, 2007/2008* 2008, Bank of Uganda: Kampala.

161. Marx, A., et al., *Meta-analysis: accuracy of rapid tests for malaria in travelers returning from endemic areas.* Annals of Internal Medicine, 2005. **142**(10): p. 836-46.

162. Kilian, A.H.D., et al., *Application of the ParaSight™-F dipstick test for malaria diagnosis in a district control program.* Acta Tropica, 1999. **72**(3): p. 281-293.

163. Guthmann, J.P., et al., *Validity, reliability and ease of use in the field of five rapid tests for the diagnosis of Plasmodium falciparum malaria in Uganda.* Trans R Soc Trop Med Hyg, 2002. **96**(3): p. 254-7.

164. Abeku, T.A., et al., *Determinants of the accuracy of rapid diagnostic tests in malaria case management: Evidence from low and moderate transmission settings in the East African highlands.* Malaria Journal, 2008. **7**.

165. Bell, D., C. Wongsrichanalai, and J.W. Barnwell, *Ensuring quality and access for malaria diagnosis: How can it be achieved?* Nature Reviews Microbiology, 2006. **4**(9 SUPPL.).

166. World Health Organisation, *Management of the child with a serious infection or severe malnutrition*, 2000.

167. Griswold, M., et al., *Analysing health care costs: a comparison of statistical methods motivated by medicare colorectal cancer charges.* Biostatistics, 2004. **1**(1): p. 24.

168. Briggs, A.H. and A.M. Gray, *Handling uncertainty in economic evaluations of healthcare interventions.* British Medical Jornal, 1999. **319**: p. 4.

169. World Health Organisation, *Indicators for IMCI at first-level facilities*, 2001, Child and adloscent health and development, World Health Organisation: Geneva.

170. CORE, et al., *Manual for rapid health facility assessment: a tool to enhance quality and access at the primary health care level*, CORE, et al., Editors.

171. USAID/CHSHP, *Kabeho mwana "the living child": rapid child health service provision assessment for quality and access at the health center level. Baseline report, march 2007*, 2007, USAID/CHSHP cooperate agreement.

172. World Health Organisation, *IMCI-MCE: Report of the IMCI Health facility Survey in Tanzania*, undated, Department of Child and Adolescent Health and Development of the World Health Organization.

173. Uganda Ministry of Health, *Health Sector Strategic Plan III, 2010/11 - 2014/15*, M.o.H. (Uganda), Editor undated.

174. Uganda Ministry of Health, *Health Sector Strategic Plan II, 2005/06 - 2009/2011; Volume 1*, M.o.H. (Uganda), Editor undated

175. Uganda Ministry of Health, *Management of Uncomplicated Malaria: A practical Guide for Health Workers*, in *Malaria Control Programme*2005.

176. Management Sciences for Health and Euro Health Group, *Managing drug supply. The Selection, Procurement, Distribution, and Use of Pharmaceuticals*. 2nd, Revised and expanded ed1997: Kumarian Press.

177. Tumwine, Y., *Availability and Expiry of Essential Medicines and Supplies During the 'Pull' and 'Push' Drug Acquisition Systems in a Rural Ugandan Hospital.* Tropical Journal of Pharmaceutical Research, 2010. **9**(6): p. 8.

178. Raja, S., S. Wilbur, and B. Blackburn, *Uganda Logistics Systems for Public Health Commodities: An Assessment Report*, 2000, John Snow Inc./Family Planning Logistics Management (FPLM), United States International Development Agency/Washington.

179. Michael Siegrist, Timothy C. Earle, and H. Gutscher, *Uncertainty and Scepticism in the Public Mind*, in *Trust in Risk Management*, Michael Siegrist, Timothy C. Earle, and H. Gutscher, Editors. 2010.

180. Fisher, R.J., *Social Desirability Bias and the Validity of Indirect Questioning.* Journal of Consumer Research, 1993. **20**(2): p. 303-315.

181. FANTA-2 Project, *Layers for Title II: An Implementation Manual. Washington, DC: Food and Nutrition Technical Assistance II Project, Academy for Educational Development, 2009.*, 2009.

182. Africa health workforce observatory, *Human Resources for Health; Country Profile: Uganda*, 2009.

183. Boone, T., R. Ganeshan, and R.L. Hicks, *Learning and Knowledge Depreciation in Professional Services.* MANAGEMENT SCIENCE, 2008. **54**(7): p. 1231-1236.

184. Ramanarayanan, S., *Does practice make perfect: an empirical analysis of learning-by-doing in cardiac surgery.* Mimeo, North Western University, 2006.

185. Ofori-Adjei, D. and D. Arhinful, *Effect of training on the clinical management of malaria by medical assistants in Ghana.* Social Science & Medicine, 1996. **42**(8): p. 1169 - 1176.

186. Hoshaw-Woodard, S., *Description and comparison of the methods of cluster sampling and lot quality assurance sampling to assess immunization coverage*, W.H. Organisation, Editor 2001.

187. Terlouw, A., et al., *Optimized age-based dosing regimens of the new fixed-dose combinations artesunate-amodiaquine and artesunate mefloquine for the treatment of uncomplicated falciparum malaria.* Tropical Medicine & International Health, 2007. **12**(Issue Suppl).

188. Jones, S.R.G., *Was There a Hawthorne Effect?* American Journal of Sociology, 1992. **98**(3): p. 451-468.

189. Rowe, S.Y., et al., *The influence of observation and setting on community health workers' practices.* International Journal for Quality in Health Care, 2006. **18**(4): p. 299-305.