



Université d'Ottawa • University of Ottawa



Université d'Ottawa • University of Ottawa

FACULTÉ DES ÉTUDES SUPÉRIEURES
ET POSTDOCTORALES

FACULTY OF GRADUATE AND
POSTDOCTORAL STUDIES

GUENTHER, Tanya

AUTEUR DE LA THÈSE - AUTHOR OF THESIS

M.Sc. (Epidemiology)

GRADE - DEGREE

Department of Epidemiology and Community Medicine

FACULTÉ, ÉCOLE, DÉPARTEMENT - FACULTY, SCHOOL, DEPARTMENT

TITRE DE LA THÈSE - TITLE OF THE THESIS

Towards Home and Community Based Approaches For Managing Childhood
Fever in Malaria Endemic Areas

G. Wells

DIRECTEUR DE LA THÈSE - THESIS SUPERVISOR

P. Tugwell

CO-DIRECTEUR DE LA THÈSE - THESIS CO-SUPERVISOR

EXAMINATEURS DE LA THÈSE - THESIS EXAMINERS

R. Nair

B. Wilson

J.-M. De Koninck, Ph.D.

LE DOYEN DE LA FACULTÉ DES ÉTUDES
SUPÉRIEURES ET POSTDOCTORALES

SIGNATURE

DEAN OF THE FACULTY OF GRADUATE
AND POSTDOCTORAL STUDIES

**TOWARDS HOME AND COMMUNITY BASED APPROACHES FOR
MANAGING CHILDHOOD FEVER IN MALARIA ENDEMIC AREAS**

By

TANYA GUENTHER

Thesis submitted to the Faculty of Graduate and Postdoctoral Studies in partial
fulfillment of the requirements for the MSc degree in Epidemiology

University of Ottawa

May 2003



National Library
of Canada

Bibliothèque nationale
du Canada

Acquisitions and
Bibliographic Services

Acquisitons et
services bibliographiques

395 Wellington Street
Ottawa ON K1A 0N4
Canada

395, rue Wellington
Ottawa ON K1A 0N4
Canada

Your file *Votre référence*
ISBN: 0-612-90076-2
Our file *Notre référence*
ISBN: 0-612-90076-2

The author has granted a non-exclusive licence allowing the National Library of Canada to reproduce, loan, distribute or sell copies of this thesis in microform, paper or electronic formats.

L'auteur a accordé une licence non exclusive permettant à la Bibliothèque nationale du Canada de reproduire, prêter, distribuer ou vendre des copies de cette thèse sous la forme de microfiche/film, de reproduction sur papier ou sur format électronique.

The author retains ownership of the copyright in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author's permission.

L'auteur conserve la propriété du droit d'auteur qui protège cette thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

In compliance with the Canadian Privacy Act some supporting forms may have been removed from this dissertation.

Conformément à la loi canadienne sur la protection de la vie privée, quelques formulaires secondaires ont été enlevés de ce manuscrit.

While these forms may be included in the document page count, their removal does not represent any loss of content from the dissertation.

Bien que ces formulaires aient inclus dans la pagination, il n'y aura aucun contenu manquant.

Canada

Abstract

Febrile illnesses associated with malaria and pneumonia account for a large proportion of child mortality in sub-Saharan Africa. Three sub-studies were conducted to contribute to the development and evaluation of programs to improve home and community management of childhood fever. The first sub-study evaluated the ability of caregivers to identify fever in their child using palpation through a systematic review. Combination of results from 12 studies demonstrated that caregivers perform well in identifying fever. The second sub-study explored the nature and determinants of caregiver fever management practices using logistic regression and community-level data from Uganda. Caregiver education level, child age and household size were associated with home treatment while distance, region and initial home treatment practices were associated with care seeking. Clustering of fever management practices was significant at the village level. The third sub-study prepared a set of proposal guidelines and sample size framework for a multicentre, cluster-randomized trial of an integrated approach.

Acknowledgements

I would like to sincerely thank all those who have contributed to this thesis.

To my main supervisor, Dr. George Wells, for his thoughtful support, guidance, encouragement and enthusiasm.

To Dr. Peter Tugwell, my co-supervisor, for his helpful comments and suggestions.

To Dr. George Pariyo at the Institute of Public Health in Uganda, for his kindness and hospitality and for generously providing access to the C-IMCI data.

To members of the HCMMP working group who provided support and opportunity.

To my family. A special thanks to my Dad who carefully read through the entire first draft and provided many helpful suggestions, to my Mom for keeping me in her prayers and to Karyn for listening.

To Karin Källander, my 'sister' in Sweden, for her amazing friendship and helpful comments.

To Michael Garner – you know who you are. Thanks for your excellent support, friendship and epidemiological editing.

To Christine Navarro, for putting up with my piles of paper and miscellaneous clutter.

A big thanks to Anne Gray for all her help and especially for her patience.

To all my climbing friends, who kept me sane.

Table of Contents

Abstract.....	i
Acknowledgements.....	ii
Table of Contents.....	iii
List of Tables.....	vi
List of Figures.....	vii
List of Acronyms.....	ix
1.0 INTRODUCTION.....	1
1.1 OVERVIEW OF THESIS.....	4
2.0 SUB-STUDY I: THE ABILITY OF CAREGIVERS TO IDENTIFY FEVER IN THEIR CHILDREN – A SYSTEMATIC REVIEW.....	7
2.1 BACKGROUND.....	7
2.1.1 Overview of fever in children.....	7
2.1.1.1 Incidence and prevalence.....	8
2.1.1.2 Sources of fever.....	9
2.1.2 Fever measurement and monitoring.....	10
2.1.2.1 Objective assessment of fever.....	10
2.1.2.2 Caregiver assessment of fever.....	11
2.1.3 Rationale for meta-analysis.....	12
2.2 OBJECTIVES.....	13
2.3 METHODS.....	13
2.3.1 Location of primary studies.....	13
2.3.2 Study selection.....	14
2.3.3 Review of study quality and applicability.....	14
2.3.4 Data extraction.....	16
2.3.5 Data synthesis and analysis.....	17
2.4 RESULTS.....	23
2.4.1 Search results.....	23
2.4.2 Study characteristics.....	24
2.4.3 Study quality.....	25
2.4.4 Diagnostic accuracy.....	27
2.5 DISCUSSION.....	35
2.5.1 Quality of primary studies.....	35
2.5.2 Interpretation of main findings.....	37
2.5.3 Factors influencing diagnostic accuracy.....	40
2.5.4 Applications for clinical and community settings.....	41
2.5.5 Limitations.....	43

3.0	SUB-STUDY II: SOCIO-DEMOGRAPHIC AND HEALTH CARE DETERMINANTS OF CAREGIVER FEVER MANAGEMENT PRACTICES.....	45
3.1	BACKGROUND.....	45
3.1.1	Overview of caregiver management practices for childhood fever.....	45
3.1.2	Factors influencing caregiver management practices for childhood fever....	48
3.1.3	Study rationale.....	50
3.2	OBJECTIVES.....	51
3.3	METHODS.....	51
3.3.1	Study area.....	51
3.3.2	Data source.....	52
3.3.3	Sampling methodology.....	53
3.3.4	Survey respondents.....	55
3.3.5	Study questionnaire.....	55
3.3.6	Definition of dependent and independent variables.....	56
3.3.6.1	Dependent variables.....	57
3.3.6.2	Independent variables.....	58
3.3.7	Data entry and quality assurance.....	59
3.3.8	Data analysis.....	59
3.3.8.1	Alternating logistic regressions.....	61
3.3.8.2	Approach to modeling.....	65
3.4	RESULTS.....	66
3.4.1	Sample characteristics.....	67
3.4.1.1	Socio-demographic and health care characteristics.....	67
3.4.1.2	Fever management practices.....	68
3.4.2	Factors associated with fever management practices.....	69
3.4.2.1	Home treatment with an antimalarial.....	70
3.4.2.2	Care seeking.....	72
3.4.2.3	Promptness of care seeking.....	74
3.4.3	Clustering of fever management practices.....	76
3.5	DISCUSSION.....	78
3.5.1	Interpretation of factors associated with caregiver fever management practices.....	78
3.5.2	Interpretation of clustering of behaviours.....	85
3.5.3	Implications for home and community programs.....	87
3.5.4	Limitations.....	88
4.0	SUB STUDY III: PROPOSAL DEVELOPMENT GUIDELINES AND SAMPLE SIZE FRAMEWORK FOR THE HOME AND COMMUNITY MANAGEMENT OF MALARIA AND PNEUMONIA (HCMMP) STUDY.....	92
4.1	BACKGROUND.....	92
4.1.1	Overview of the HCMMP study.....	92
4.1.2	Development of proposal development guidelines and sample size framework for the HCMMP study.....	95

4.2	OBJECTIVES.....	96
4.3	METHODS.....	96
4.3.1	Development of the HCMMP proposal development guidelines.....	96
4.3.2	Framework for sample size and power.....	99
4.4	RESULTS.....	103
4.4.1	Proposal development guidelines.....	103
4.4.2	Sample size and power sensitivity analysis.....	103
4.4.2.1	Estimation of sample size parameters.....	103
4.4.2.2	Sensitivity analysis.....	105
4.5	DISCUSSION.....	111
4.5.1	Proposal development guidelines.....	112
4.5.2	Sample size framework.....	116
4.5.3	Next steps.....	119
5.0	CONCLUSIONS.....	121
6.0	REFERENCE LIST.....	124
7.0	APPENDICES.....	136
	APPENDIX IA: SEARCH STRATEGY.....	136
	APPENDIX IB: EXCLUDED STUDIES.....	137
	APPENDIX IIA: SELECTED QUESTIONS FROM THE COMMUNITY INTEGRATED MANAGEMENT OF CHILDHOOD ILLNESS (C-IMCI) HOUSEHOLD SURVEY	141
	APPENDIX IIIA: SUMMARY OF ENTEBBE MEETING RESULTS	144
	APPENDIX IIIB: HCMMP PROPOSAL DEVELOPMENT GUIDELINES.....	145

List of Tables

Sub-Study I

Table 2.1	Study selection criteria.....	14
Table 2.2	Characteristics of included studies.....	25
Table 2.3	Results of study quality evaluation.....	26
Table 2.4	Summary of diagnostic information for included studies.....	27
Table 2.5	Regression model parameters (weighted and unweighted).....	30
Table 2.6	Regression model parameters for the subgroup analysis.....	33

Sub-Study II

Table 3.1	Description of sampling units for the Community Integrated Management of Childhood Illness (C-IMCI) survey.....	54
Table 3.2	Definition and coding of independent and dependent variables.....	56
Table 3.3	Home care practices for childhood fever before or instead of seeking care reported by caregivers.....	68
Table 3.4	Summary of home care and care seeking practices for childhood fever reported by caregivers.....	69
Table 3.5	Number of children per village subcluster and health centre cluster by model type.....	70
Table 3.6	Unadjusted and adjusted odds ratios (OR) and 95% confidence intervals for home treatment with an antimalarial among caregivers reporting a child with fever during the last two weeks.....	71
Table 3.7	Unadjusted and adjusted odds ratios (OR) and 95% confidence intervals for seeking outside care among caregivers reporting a child with fever during the last two weeks.....	73
Table 3.8	Unadjusted and adjusted odds ratios (OR) and 95% confidence intervals for promptness of care seeking among caregivers reporting a child with fever during the last two weeks.....	75
Table 3.9	Unadjusted and adjusted within-village and within-health centre catchment area pairwise odds ratios (PWOR) for home treatment, care seeking and promptness of care seeking	77

Sub-Study III

Table 4.1	Total and under-five population estimates for selected countries by administrative level.....	105
Table 4.2	Sample size requirements for a two year, three-arm cluster randomized trial to reduce under-five mortality in the malaria arm by 25% compared to the control arm, assuming a baseline mortality rate of 80/1,000 live births and a coefficient of variation of 0.25.....	106

List of Figures

Figure 1.1	Overview of thesis components in relation to the pathway to child survival for management of childhood fever.....	4
Sub-Study I		
Figure 2.1	Flow diagram of search strategy results.....	23
Figure 2.2	Plot of sensitivity and specificity with 95% confidence intervals and the overall pooled estimate using the random effects model.....	28
Figure 2.3	Plot of D versus S with the unweighted ordinary least squares regression line for the all studies (dashed line) and outliers removed (solid line) analyses.....	29
Figure 2.4	Plot of sensitivity against 1-specificity with 95% confidence intervals and the unweighted (upper) and weighted (lower) summary receiver operating characteristic (SROC) curves for the all studies and outliers removed analyses.....	31
Sub-Study II		
Figure 3.1	Map of Uganda showing districts included in the Community Integrated Management of Illness (C-IMCI) survey.....	53
Figure 3.2	Flow diagram of Community Integrated Management of Illness (C-IMCI) survey respondents.....	67
Figure 3.3	Unadjusted and adjusted within-village and within-health centre catchment area pairwise odds ratios (PWORs) for home treatment with an antimalarial, care seeking, and promptness of care seeking....	77
Sub-Study III		
Figure 4.1	Total number of children and number of clusters per study arm by cluster size assuming baseline mortality rate of 80/1,000 child-years, a coefficient of variation of 0.25, an effect size of 0.25, an alpha of 0.05 and study power of 0.80.....	107
Figure 4.2	Number of clusters per arm by cluster size across selected parameters assuming a mid-range set of parameters as appropriate (a coefficient of variation of 0.25, an effect size of 0.25, a baseline mortality rate of 80/1,000 child-years), an alpha of 0.05 and study power of 0.80.....	108
Figure 4.3	Study power by number of clusters per study arm and cluster size (CS) assuming a baseline mortality rate of 80/1,000 child-years, a coefficient of variation of 0.25, an effect size of 0.25 and an alpha of 0.05.....	109
Figure 4.4	Study power by cluster size and by coefficient of variation assuming a baseline mortality rate of 80/1,000 child-years, ten clusters per study arm, an effect size of 0.25 and an alpha of 0.05.....	110

Figure 4.5 Number of clusters per study arm by cluster size and loss to follow-up assuming a coefficient of variation of 0.25, an effect size of 0.25, and an alpha of 0.05..... 110

List of Acronyms

ALR	Alternating Logistic Regressions
ARI	Acute Respiratory Infection
BMR	Baseline Mortality Rate
C-IMCI	Community Integrated Management of Childhood Illness
CI	Confidence Interval
CIHR	Canadian Institutes for Health Research
CONSORT	Consolidated Standard of Reporting Trials
CS	Cluster size
CV	Coefficient of Variation
DHS	Demographic and Health Survey
DOR	Diagnostic Odds Ratio
ES	Effect Size
FN	False Negative
FP	False Positive
FPR	False Positive Rate
GEE	Generalized Estimating Equations
GLM	Generalized Linear Model
HCMMP	Home and Community Management of Malaria and Pneumonia
HCW	Health Care Worker
HIV	Human Immunodeficiency Virus
ICC	Intracluster Correlation Coefficient
IEC	Information, Education, Communication
IF	Inflation Factor
IMCI	Integrated Management of Childhood Illness
KM	Kilometre
MRC	Medical Research Council
OLR	Ordinary Logistic Regression
OR	Odds Ratio
PWOR	Pair Wise Odds Ratio
RBM	Roll Back Malaria
REM	Random Effects Model
ROC	Receiver Operating Characteristic
SDOR	Summary Diagnostic Odds Ratio
SE	Standard Error
SROC	Summary Receiver Operating Characteristic
SSA	Sub-Saharan Africa
STARD	Standards for Reporting of Diagnostic Accuracy
TDR	Special Programme for Tropical Diseases Research
TN	True Negative
TP	True Positive
TPR	True Positive Rate
UDHS	Uganda Demographic Health Survey
WHO	World Health Organization

1.0 INTRODUCTION

In sub-Saharan Africa (SSA) child mortality rates remain the highest in the world, with an average under-five mortality rate of 175/1,000 live births¹. While under-five mortality rates declined 11% globally between 1990 and 2000, progress was minimal to non-existent in SSA and it is anticipated that by 2010 half of all under-five deaths worldwide will occur in this region¹. Malaria and acute respiratory infections (ARI), of which most are pneumonia, are leading causes of these deaths and each account for upwards of one-fifth of overall under-five mortality in SSA²⁻⁴. The majority of deaths due to malaria and pneumonia result from lack of early diagnosis and prompt appropriate treatment with an effective antimalarial and/or antibiotic^{4,5}. More than 80% of child deaths occur in the home without any contact with a trained health provider or a health facility^{6,7}. The reasons for this are many and include prevailing traditional beliefs, cultural habits, poor education, poverty, low quality services and the rapidity with which severe complications and death can follow initial symptoms, often within as little as 24 to 48 hours of symptom onset⁸.

Mortality and morbidity control strategies for both malaria and pneumonia are based on presumptive management^{4,9}. This approach makes a diagnosis based on simple signs and symptoms selected for their predictive value and easily identified by health workers with minimal training and in the absence of laboratory support⁹. For example, in malaria endemic areas, fever or history of fever is treated as malaria⁹. Over the past few years, efforts have been undertaken to integrate the management of common childhood infections as part of the Integrated Management of Childhood Illnesses (IMCI) initiative. The IMCI recognizes that sick children often suffer from more than one

condition concurrently and that focussing on a single disease may overlook such children. This is a key concern for malaria and pneumonia in areas where both diseases are endemic. Fever is a common initial sign of both malaria and pneumonia and signs of disease severity also overlap (inability to eat, convulsions and difficulty breathing)¹⁰. As a result, these conditions cannot be distinguished without considerable laboratory support, even by a trained health worker^{10;11}. Several studies indicate that approximately one-third of children presenting to health clinics in sub-Saharan Africa satisfy both the malaria and pneumonia presumptive case definitions and require dual treatment¹⁰⁻¹².

To date, the focus of implementation of presumptive management under the direction of the IMCI has been largely limited to the health facility level. The impact has been predictably low in SSA where access to even basic health care, particularly in rural areas, is generally poor^{13;14}. Most illness management occurs within the home with little involvement of the health facility and the recognition of illness and provision of care by a child's caregiver is the most important contributor to a child's health and survival¹⁵. Thus, effective case management in the home becomes even more critical and caregivers must be able to recognize and interpret illness onset and promptly provide care or seek care at an appropriate source.

Efforts to extend the IMCI approach and to focus on care provided in the home and community are just beginning, and strategies supported by an appropriate evidence base are urgently needed¹⁶. One promising approach undergoing development is a set of home and community based programs that apply the principles of presumptive treatment and integrated management of childhood fever closer to the home and community. This approach recognizes that the achievement of prompt and appropriate treatment of fever,

necessary to reduce morbidity associated with malaria and other infectious diseases in young children, depends in large part on behaviours initiated by their caregivers, particularly in rural areas where access to formal health facilities is often limited. These strategies are designed to address the complex array of factors contributing to child mortality and integrate the involvement of a wide range of stakeholders (caregivers, drug vendors, traditional healers and formal medical staff) who play a role in the management of these illnesses. They aim both to improve caregiver awareness of illness in the home and make essential medicines available as close to the home as possible through trained, resident community based agents.

The general purpose of this thesis is to contribute to the development and evaluation of programs to improve home and community management of common childhood illnesses, focussing on the management of childhood fever by caregivers in rural areas of sub-Saharan Africa where malaria is endemic. Specific aims are:

- i. to evaluate the ability of caregivers to accurately identify fever in their child using subjective assessment;
- ii. to explore the nature and determinants of caregiver responses to childhood fever; and
- iii. to develop initial design requirements for a study of an integrated approach to the home and community management of childhood fevers.

These study aims and their integration are illustrated in Figure 1.1.

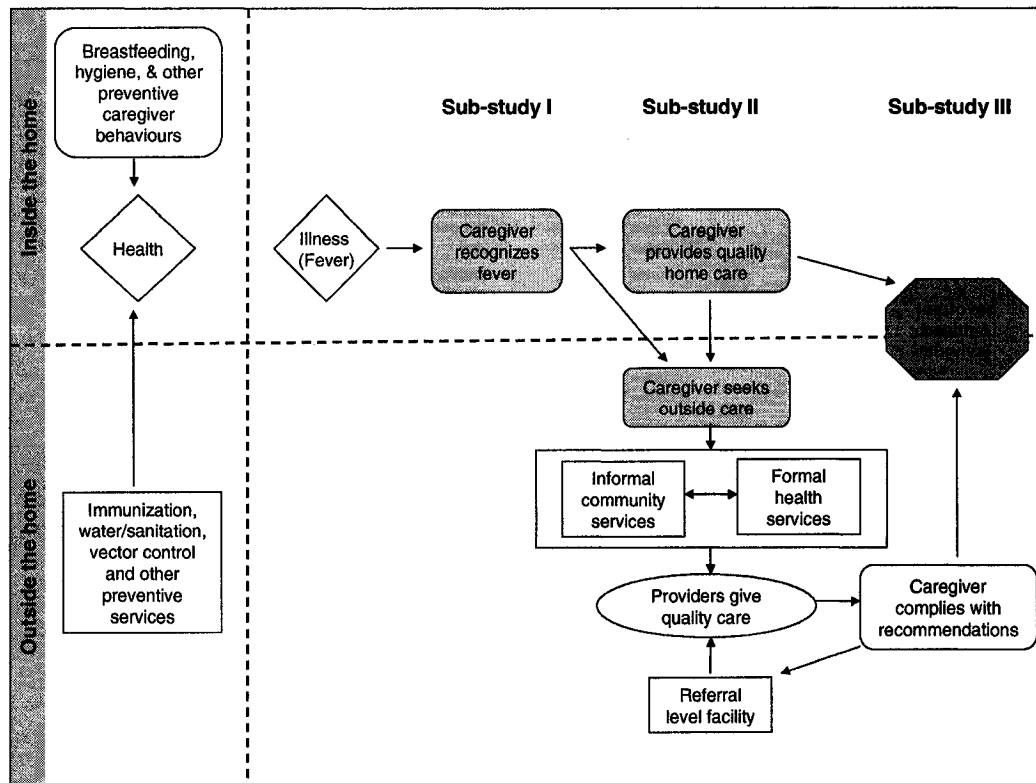


Figure 1.1. Overview of thesis components in relation to the pathway to child survival for management of childhood fever. Adapted from the 'The Pathway to Survival' published in "Overcoming Remaining Barriers: The Pathway to Survival" by Ronald Waldman, Alfred Bartlett, Carlos Campbell and Richard Steketee¹⁵. Shaded areas denote areas explored in this thesis.

1.1 OVERVIEW OF THESIS ORGANIZATION

This thesis is comprised of three sub-studies. Each sub-study is a stand-alone document that contributes to the overall theme of home and community management of fever in areas where malaria and pneumonia co-exist. These sub-studies support the development and design of an ongoing multi-country study, the Home and Community Management of Malaria and Pneumonia (HCMMP) study, which aims to evaluate an integrated approach to fever management.

The first sub-study is a systematic review and meta-analysis of the ability of caregivers, who are primarily mothers, to subjectively identify fever in their young

children. Fever is often the earliest sign that an illness is present and the ability of caregivers to detect it is an important prerequisite for initiating care in the home and seeking appropriate care outside the home¹⁷. Recently developed meta-analysis techniques for diagnostic studies are used to combine studies to explore overall diagnostic accuracy as well as factors that influence accuracy. These results are then applied to clinic and community settings where the diagnosis of fever relies largely on subjective caregiver assessments.

The second sub-study is a primary analysis of a dataset collected at the community level as part of an ongoing evaluation of the community component of the IMCI (C-IMCI) in Uganda. The achievement of prompt and appropriate treatment of fever depends in large part on the home treatment and care seeking behaviours initiated by their caregivers. This analysis explores the influence of socio-demographic and health care factors and whether these behaviours cluster in communities and health unit catchment areas. Three aspects of caregiver management practices for childhood fever are examined: home treatment with an antimalarial before or instead of seeking care; seeking external care; and promptness of care seeking. Understanding the dynamics of and what factors influence caregiver responses to fever can help focus strategies to promote optimal behaviour and improve case management of fever.

The final component of this thesis presents initial work on the design of the HCMMP study. The HCMMP study is a multi-country, cluster-randomized controlled trial to compare the safety and effectiveness of three methods of providing care for childhood febrile illness – one that provides home and community based care for malaria and pneumonia together, one that provides home and community based care for malaria

alone, and one that is representative of normal care and based primarily at the health facility. Two primary outputs prepared were a set of proposal development guidelines to develop a common protocol framework and a framework for assessing sample size and power requirements to account for the cluster nature of the study. This work was undertaken in collaboration with an international working group as well as potential country level researchers.

These three sub-studies address issues related to the presumptive management of childhood illness at the home and community levels in sub-Saharan Africa. Each sub-study is self-contained and presented under the following section headings: background, objectives, methods, results and discussion. These sections are followed by overall conclusions. Associated reference lists and appendices for all sub-studies are placed at the end.

2.0 SUB-STUDY I: THE ABILITY OF CAREGIVERS TO IDENTIFY FEVER IN THEIR CHILDREN – A SYSTEMATIC REVIEW

2.1 BACKGROUND

Caregiver recognition of illness is the first and most critical step in helping a sick child receive prompt and appropriate management¹⁵. Fever is often the earliest sign that an illness is present and the ability of caregivers to detect it is an important prerequisite for initiating care in the home and seeking appropriate care outside the home^{15;17}.

Identifying the presence of fever is critical in areas where febrile illnesses such as malaria and pneumonia are endemic as these diseases remain major causes of child mortality and morbidity and require prompt attention^{1;4;18}. In many areas of the world, and particularly those where febrile illnesses account for a large proportion of child mortality, caregiver detection of fever is subjective, primarily by palpation of the skin. This sub-study undertakes a systematic review to evaluate the ability of caregivers to identify fever in their young children and explore factors that influence this ability. The applications for clinical and community settings in poor countries are also reviewed.

2.1.1 Overview of fever in children

Fever is defined as a carefully controlled upward shift in regulated body temperature¹⁷. The febrile response is initiated in response to microorganisms, immune complexes, or other sources of inflammation in the body and therefore is a common initial symptom for a wide variety of illnesses in children^{15;19;20}. Children develop the ability to mount a febrile response during the first few months of life and feverish temperatures are uncommon among those less than two months of age^{17;19}. Fever itself is rarely harmful; rather a moderately elevated rise in core temperature generally improves the ability of the immune system to fight infection by accelerating its response to invasion

and inhibiting the growth rate of several key pathogens, including the parasitological agent responsible for malaria, *Plasmodium*¹⁹. A small proportion of children will experience febrile convulsions, but these usually occur early in the course of fever and resolve without consequence¹⁹.

2.1.1.1 Incidence and prevalence

Only a few studies provide fever incidence estimates determined by thermometer readings and based on frequent surveillance over a sufficient time period. These studies indicate that fever occurs frequently and incidence depends on many factors, including age profile of children studied and level of malaria transmission. In a rural Senegalese village where malaria is highly endemic, 94 children followed daily for four months experienced a mean number of fever episodes during this period of 2.9 for children aged 1 to 23 months, 2.5 for those between 2 and 6 years, and 0.8 for those aged 7 to 12²¹. In Mali, the nine-day cumulative incidence of fever was measured in over 800 children aged 1 to 12 years during the rainy (high malaria) season and the dry (low malaria) season. The prevalence of fever on the first day of study was 6.2% in the dry season and 12.8% in the rainy season. Over the following nine-day period, 2% developed fever during the dry season and 8.2% in the rainy season. Younger children aged 1 to 3 years were more than twice as likely as older children to develop fever in each season²².

The prevalence of fever in children measured in communities exhibits considerable variation, with estimates ranging from <1% to 38% and is strongly related to both age and malaria transmission. Fever prevalence follows a consistent age-related pattern, being highest among children between three months and three years, and then declining throughout childhood²²⁻²⁶. The relation of fever prevalence to malaria transmission was

evident in numerous studies and very high levels of fever have been recorded in areas where malaria is highly endemic. A community surveillance study in a cohort of 1,500 children under five years old in northern Ghana reported a point prevalence of fever ranging from 2 to 8% in the low malaria season to between 6 and 14% in the high malaria season²³. Among children aged 0 to 14 in the Ivory Coast, where the average inoculation rate is 400 infective bites per year, overall prevalence of fever was 21%, and estimates reached 38% in rural areas just after a rainy season when malaria transmission is most intense²⁷. Similarly, in a rural area of the Republic of Guinea, the prevalence of fever during the rainy season was 23% in a sample of 784 children less than five years of age²⁸. Conversely, low fever prevalence was reported among 317 children living in an area of Kenya with relatively low malaria transmission (6%) and among 799 under-five children in Nepal with no malaria transmission(0.6%)^{29;30}.

2.1.1.2 Sources of fever

Common sources of fever in children include acute respiratory infections (ARI), dysentery, meningitis, measles, ear infections, sore throat, and depending on the area, malaria, dengue hemorrhagic fever and symptomatic human immunodeficiency virus (HIV) infection³¹. In industrialized countries, self-limiting viral infections and/or otitis media account for most fevers. Less than 10% of fevers seen in emergency rooms are considered serious and most of these are due to pneumonia¹⁹. In low-income countries, potentially serious causes and consequences of fever are much more common due to elevated exposure to more dangerous disease agents, including that which causes malaria, as well as compromised immune systems resulting from poor nutrition. In sub-Saharan Africa, community based surveillance studies indicate that malaria is responsible for

between 20% to 60% of fevers observed in young children (Tanzania: \approx 20%; Benin: 33%; The Gambia: 40%; Senegal: 60%)^{24;26;32;33}. At health facilities, the proportion of fever in children associated with malaria is even higher. In Uganda, 64% of feverish children under-five years were positive for malaria parasites³⁴. Similarly, approximately 68% of fevers children aged 0 to 9 years in a rural Tanzanian health clinic were due to malaria according to microscopy of blood smears³⁵. The remainder were due to influenza, pneumonia, common cold, tooth eruptions or unspecified³⁵.

2.1.2 Fever measurement and monitoring

2.1.2.1 *Objective assessment of fever*

Four main types of clinical thermometers are used to measure body temperature in children: liquid (mercury) in glass thermometers; digital thermometers; liquid crystal thermometers; and radiometers for the ear canal³⁶. Of these, mercury and digital thermometers are the most widely used in outpatient and emergency clinics and by parents due to their low cost and availability³⁶. Liquid crystal thermometers are not recommended due to their low accuracy and poor reliability³⁶. A child's body temperature can be measured at a number of sites, the most common being the mouth, rectum, axilla (armpit), and tympanic membrane (ear). Rectal temperatures equal to or exceeding 38.0°C are considered febrile¹⁷. Axillary, oral and tympanic temperatures are usually slightly lower than rectal temperatures and are considered feverish at and above, 37.0°C, 37.5°C and 37.9°C (37.6°C for children 11 and older), respectively³⁶.

The gold standard for measurement of body temperature and assessment of fever in children is rectal temperature as it most closely approximates core body temperature³⁷. Tympanic temperatures provide maximum readings within a few seconds and require

minimal patient contact, but are not recommended in paediatric populations due to their poor correlation with core body temperature¹⁹. Axillary temperature readings also correlate poorly with core body temperature and with concurrent rectal measurements¹⁹. They also require a long time to achieve optimal measurement; less than one-fifth of axillary temperatures reach their maximum within five minutes compared to 90% of rectal temperatures within four minutes¹⁹. Despite their limitations, axilla measurements are preferred in paediatric departments due to their non-invasive nature and lower risk of injury and cross infection^{36;37}.

2.1.2.2 Caregiver assessment of fever

Caregivers of young children play a critical role in the identification of fever. In the home, caregivers are in a unique position to regularly monitor their child's body temperature and recognize when he or she becomes ill and feverish. Subjective assessment, generally through palpation of the child's forehead, face/neck, trunk, or abdomen, remains the main method of fever detection for most caregivers^{38;39}. This is true even where thermometers are available, as in most industrialized countries, since many mothers prefer to rely on their own subjective measurements^{40;41}. Outside the home, caregiver assessments of fever are also widely relied on. In countries with limited infrastructure, health workers may not have access to a thermometer^{9;39}. In such situations, the assessment of body temperature will often rely solely on a caregiver report of fever. In addition, fever associated with certain serious and common febrile illnesses such as malaria can be erratic. Relying on temperature at time of consultation may not be sensitive enough and a history of fever is necessary^{9;42}. In fact, the IMCI instructs health

workers to rely on a caregiver report of fever in their assessment of a sick child even when temperature can be measured⁹.

2.1.3 Rationale for a meta-analysis

Considerable controversy surrounds the ability of caregivers to correctly identify the presence or absence fever in response to several studies reporting conflicting results^{30;38;43}. The diagnostic accuracy of caregiver assessment has implications for health care workers who must rely on a caregivers' report of fever and for emerging home and community based programs where caregivers are responsible for the initial decision to treat and case management^{44;45}. Thus, it is important to systematically evaluate the ability of caregivers to assess fever in their child and to identify factors that influence this ability. The value of meta-analysis to combine results of similar studies is well recognized for estimates of efficacy and effectiveness of health interventions. The recent development of statistical techniques and related software for combining diagnostic studies now allow these results to be pooled in a way that addresses the complexities associated with estimates of diagnostic accuracy^{46;47}. Importantly for the study of caregivers' ability to recognize fever, these techniques allow studies with different threshold values to be combined without requiring knowledge of the actual values. This approach is particularly valuable for exploring variation among studies and coming to some understanding of factors that account for this variation.

2.2 OBJECTIVES

The primary objective of this sub-study is to assess the diagnostic accuracy of caregivers' subjective assessment of fever in their child compared against fever measured with a thermometer. The secondary objective is to examine whether the accuracy of caregiver assessment differs across child and caregiver attributes and study characteristics. These objectives are undertaken for application in settings where the objective measurement of fever is rarely achievable.

2.3 METHODS

The methods followed for this review are those recommended by the Cochrane Methods Working Group on Systematic Review of Screening and Diagnostic Tests⁴⁸ as well as a series of published guidelines (Irwig et al, 1994^{49,50}; Midgette et al, 1993⁵¹; and the Australian National Health and Medical Research Council guidelines⁵²).

2.3.1 Location of primary studies

Eligible studies published by October 2002 were located through a comprehensive search of electronic bibliographic databases and reference lists from primary articles. The databases MEDLINE (1966-2002), HealthSTAR (1987-2002), CINAHL (1982-2002), Current Contents (1993 week 26 – 2002), and Dissertations Abstracts (1861-2002) were searched on OVID using a search strategy designed to identify studies addressing the three main areas under study: body temperature/fever, caregiver assessment and diagnostic accuracy. A full description of the search strategy is provided in Appendix IA. Expert search terms for diagnostic accuracy were obtained from the Evidence Based Medicine website at the University of Illinois that provides a list of filters

(<http://www.uic.edu/depts/lib/lhsp/resources/filters.shtml>). Reference lists of primary studies were searched for any relevant published or unpublished studies.

2.3.2 Study selection

One reviewer (TG) examined the titles and available abstracts of citations for potentially eligible studies. Review articles, commentaries, and unrelated studies were screened out and candidate articles were retrieved for further assessment. Each eligible study was assessed by one reviewer (TG) using the inclusion and exclusion criteria outlined in Table 2.1; only studies fulfilling all criteria were included.

Table 2.1 Study selection criteria

Selection criteria	Inclusion criteria	Exclusion criteria
Population	<ul style="list-style-type: none"> ▪ Caregivers of children 	<ul style="list-style-type: none"> ▪ Health care workers
Interventions	<ul style="list-style-type: none"> ▪ Subjective assessment by caregiver to detect current fever in child compared to an acceptable reference standard in the same child within a suitable time frame ▪ Temperature taken with a mercury, electronic or digital thermometer at rectal, oral, tympanic or axillary sites as reference standard 	<ul style="list-style-type: none"> ▪ Self-palpation to diagnose fever ▪ Volunteered history or presence of fever only
Outcome	<ul style="list-style-type: none"> ▪ Presence of fever confirmed by an acceptable reference standard ▪ Adequate information to calculate sensitivity and specificity 	
Study Design	<ul style="list-style-type: none"> ▪ Non-randomized and randomized studies comparing caregiver subjective assessment with a reference standard 	

2.3.3 Review of study quality and applicability

Studies meeting the selection criteria were evaluated by the primary reviewer (TG) using quality assessment criteria related to study design and diagnostic accuracy. In an ideal study to determine the diagnostic accuracy of a test, data are collected prospectively

in a group of consecutive patients who are representative of the target population⁵³. All patients given the test are also assessed with the gold standard and both sets of measurements are performed in a blinded fashion. Primary studies were classified according to quality characteristics identified as important in the literature on diagnostic accuracy and meta-analyses^{49;53;54}, which included:

- *Data collection.* Studies were classified as prospective (participants identified and data collected before caregiver and reference standard assessments performed), retrospective (participants and/or data collection initiated after either caregiver or reference standard assessment), or unknown⁵³.
- *Selection of subjects.* Studies are susceptible to selection bias if patients are not recruited consecutively or through appropriate random selection procedures⁵⁴. Studies were classified as random/consecutive or non-consecutive/not reported.
- *Verification of test results.* Verification bias may occur if the decision to perform the reference test is contingent upon the results of the test under study⁵⁴. Potential for verification bias was considered present if only a portion of children assessed by their caregiver were examined using an appropriate reference standard and absent if all children were also examined with the reference standard.
- *Independence of measurements.* Studies were considered blinded if the caregiver and/or the individual reading the child's temperature were unaware of the results of the other. Studies in which the test and reference assessment were performed simultaneously or where insufficient details were reported were classified as unblinded.

- *Reference details.* Studies were classified as sufficient if they reported: site of temperature measurement; type of thermometer; amount of time thermometer in place for; an indication of the time elapsed between objective and subjective assessments; and the definition of fever.
- *Test details.* Studies were classified as sufficient or non-sufficient according to the completeness of their description of how caregivers assessed the presence of fever. A description of the question asked caregivers and the technique used (i.e. anatomical site) was considered adequate information.

No overall quality score was assigned since each source of bias is likely to have a different effect, sometimes in opposing directions, and these important individual effects may be masked if they are considered simultaneously⁵⁰.

Study applicability was assessed by examining characteristics of the study setting and study population and how the caregiver assessments were performed. Information collected included the presenting health problems and prevalence of fever in the children under study, the extent of malaria in the region, the education level of caregivers, the type of caregiver (mother, father, other), the age and sex of children, and whether the study was conducted at a clinic or community level.

2.3.4 Data extraction

Diagnostic information abstracted from each study included point estimates for sensitivity and specificity and the two-by-two table with the number of true positives (TP), false negatives (FN), true negatives (TN) and false positives (FP). Information on study setting, caregiver and child characteristics, test and reference standard methodology, and the prevalence of fever were also obtained. Data were entered into

Microsoft Excel[®] spreadsheets created to record data on study characteristics, quality and diagnostic information. Two reviewers (TG and KK) completed this process independently and any disagreements were resolved by consensus.

2.3.5 Data synthesis and analysis

Summary of diagnostic accuracy features: Estimates of diagnostic accuracy and corresponding 95% confidence intervals were calculated for each primary study⁴⁸.

Sensitivity, representing the probability that the caregiver assessment is positive for fever among those children who truly have fever, was calculated as $TP/(TP+FN)$ ⁵⁵. Specificity, representing the probability that the caregiver assessment is negative for fever among those children who truly do not have fever, was calculated as $TN/(TN+FP)$ ⁵⁵. The diagnostic odds ratio (DOR) was calculated as $sensitivity/(1-sensitivity)/(1-specificity)/specificity$ ⁵³. In this case, the DOR represents the odds that a child with fever will be identified as feverish by their caregiver relative to the odds that a child without fever will be identified as feverish⁵³. The DOR is a useful value for comparison across studies because it combines sensitivity and specificity information into a single value.

Data synthesis and preparation of the SROC curve: A key component of meta-analysis is the combination of results from eligible studies to provide an overall summary estimate of effect. Sensitivity and specificity values, however, are dependent on the threshold used by caregivers and the threshold used to identify fever in the reference measurement and consequently there is often a trade-off relationship between the two that varies across studies⁴⁶. Other factors that can influence this trade-off effect include differences in the underlying populations in terms of disease spectrum and patient characteristics and study methodology⁴⁷. The existence of such a trade-off relationship

means that the sensitivity and specificity values cannot be considered separately across studies. Consequently, in meta-analysis of diagnostic studies, the pooling of results to create single summary measures for sensitivity and specificity is invalid if the true positive rate (TPR) (sensitivity) and false positive rate (FPR) (1-specificity) are positively correlated, as is commonly the case⁵¹. The degree of correlation is a measure of the trade-off relationship.

In the presence of a statistically significant positive correlation, a summary receiver operating characteristic (SROC) curve should be constructed⁵¹. A standard ROC curve for a single study plots sensitivity against the 1-specificity across a range of test cut-off values and permits observation of the trade off between sensitivity and specificity⁵⁵. A curve situated at the upper left hand corner of the graph indicates a good 'test', possessing simultaneously high sensitivity and high specificity⁵⁵. The area under the ROC curve provides an overall measure of the accuracy of the test and represents the probability that a subject with the condition chosen at random is correctly identified as having the condition compared with a randomly chosen subject without the condition⁵⁴. A value of 0.5 indicates the test is no better than chance and 1.0 represents a perfect test. A summary ROC curve plots the sensitivity and specificity results for each primary study as a single point on the curve. The summary ROC curve method is the preferred approach for combining dichotomous data from primary studies with varying thresholds and has been validated for meta-analyses of a small number of studies (about 10)^{48,56}.

The presence of a trade-off relationship in the current data was examined by observing the plot of TPR versus FPR and tested statistically using Spearman's non-parametric test of correlation⁵¹. The SROC curve was created following the methods

outlined by Moses et al⁴⁶. In this method, the TPR and FPR from each primary study are first converted to logit form so that a straight regression line can be fitted. As a consequence of measurement error, different results are obtained under ordinary least squares regression depending on whether or not logit(TPR) or logit(FPR) is chosen as the dependent variable, even though there should theoretically be no difference⁵⁰. To overcome this, the difference of the logits of a positive test for individuals with and without disease (D) is regressed on the sum of logits (S) using the conventional least squares method. The intercept and slope of the fitted line are calculated and then transformed back into ROC space to generate the SROC curve. These procedures are represented mathematically by the following series of equations:

$$\text{Logit TPR} = \log[\text{TPR}/(1-\text{TPR})]$$

$$\text{Logit FPR} = \log[\text{FPR}/(1-\text{FPR})]$$

$$D = \text{logit TPR} - \text{logit FPR}$$

$$S = \text{logit TPR} + \text{logit FPR}$$

$$\text{Regression equation: } D = \alpha + \beta S$$

Where: α = intercept term; β = regression coefficient for S

$$\text{Back-transformation equation: } \text{TPR} = \left[1 + e^{-\alpha/(1+\beta)} * (1-\text{FPR}/\text{FPR})^{(1+\beta)/(1-\beta)} \right]^{-1}$$

To address the possibility of zero cells, a continuity correction was performed by adding 0.5 to each cell in the two-by-two table (i.e. to the TP, FP, TN, and FN values), including those with no zeros⁴⁶. This correction shifts the SROC curve downwards from the upper left hand corner slightly, but protects against undefined equations⁴⁶.

The value D corresponds to the log diagnostic odds ratio and increased values indicate improved discriminatory ability of caregivers to identify fever⁴⁶. The value S

represents the threshold used by caregivers to classify a child as feverish: a positive value indicates the threshold favours sensitivity over specificity and vice versa while a value of 0 indicates neutrality between sensitivity and specificity⁵⁰. The extent to which D varies across individual studies due to variation in the test threshold is estimated by the value of the regression coefficient for S (β)⁵³. This can be assessed by observing the symmetry of the SROC curve and the magnitude and statistical significance of β . The SROC curve will be symmetric if the odds ratio is independent of the test threshold and asymmetric if the odds ratio is dependent on the test threshold⁵⁰. In general, parameter values for S that lie between -0.5 and 0.5 will produce a symmetric curve⁴⁷. When β for S is close to 0 and not statistically significant, the model intercept (α) is an estimate of the common log diagnostic odds ratio across studies, termed the summary diagnostic odds ratio (SDOR)⁴⁶.

In studies of diagnostic accuracy, the SDOR is an advantageous summary value since it stays constant across the range of sensitivity and specificity values and therefore provides a single measure of diagnostic accuracy across studies⁵⁰. The SDOR can be estimated directly from the regression equation by transforming the intercept into an odds ratio ($\exp(\alpha)$) or by combining the DORs from the primary studies using regular meta-analysis methods for pooling odds ratios. The DerSimonian and Laird random effects methods are recommended over Mantel-Haenszel for combining odds ratios since the latter is unable to account for interstudy variation^{46;57}. For the current analysis, the SDOR was derived directly from the regression equation as this approach provides the most conservative estimate and is simple to obtain⁴⁹.

As recommended by the Cochrane Methods Working Group, both weighted and unweighted regression analyses were performed⁴⁸. For the weighted analysis, results of

each primary study were weighted by the inverse of the variance of the log odds ratio (Weight = $[1/(1/TP+0.5) + (1/FN+0.5) + (1/FP+0.5) + (1/TN+0.5)]$). A weighted approach assumes that there is no between-study variation and results are susceptible to being determined by one or two large studies⁴⁶. In contrast, an unweighted analysis weights each study equally and thus accounts for between-study variation, while essentially ignoring any within-study variation⁴⁷. The unweighted regression analysis is preferred over the weighted analysis since in meta-analysis, between-study variability is more important to account for than within-study variability⁴⁹.

Primary studies with unusually high or low FPR and TPR values can produce estimates for D and S that disproportionately affect the slope of the regression line⁴⁶. To address this, two series of analysis were performed. The first included all studies while the second excluded studies considered outliers based on extreme TPR or FPR values⁴⁶.

Obtaining an estimate of overall accuracy: The area under the ROC curve, which is commonly used in single diagnostic studies to estimate overall test accuracy, is not appropriate in a meta-analysis of diagnostic studies because the SROC curve is only plotted over the range of included studies⁴⁶. Two alternate approaches, both based on the SROC curve, were undertaken. The first approach, suggested by Moses et al, identifies the value on the SROC curve where sensitivity equals specificity⁴⁶. This common value (A) identifies the position of the SROC curve in relation to the desirable upper left hand corner where sensitivity and specificity both equal 1.0; the closer A is to 1.0, the greater the test accuracy. The A and its standard error (SE) are calculated from the SROC regression equation as follows:

$$A = [1 + e^{-a/2}]^{-1}$$

$$SE(A) = [SE(\alpha)/8(\cosh(\alpha/4)^2)].$$

The second approach derives summary estimates of sensitivity and specificity by identifying a plausible value for either parameter (i.e. the pooled estimate, setting specific data) and locating the corresponding value on the SROC curve for the remaining parameter⁵⁰. This approach takes into account the trade-off relationship and allows values to be tailored to a given situation using available data.

Subgroup analysis: The regression methods employed to derive the SROC curve were used to explore the influence of study and patient factors on variability in diagnostic accuracy across studies⁵⁰. The basic regression equation was expanded to include covariates: $D = \alpha + \beta S + \theta X$, where X represents a covariate and θ its parameter estimate. Variables were considered for inclusion if they were identified in the literature as potentially important and had a minimum of at least three studies in each subgroup^{53,54}. Variables were entered separately into the main SROC models (weighted and unweighted) and the regression coefficients were examined to evaluate the impact of the characteristic on diagnostic accuracy⁴⁹.

Statistical Software: Meta-Test[®] (Joseph Lau, New England Medical Center, Massachusetts) was used to prepare the SROC curves and to calculate exact binomial 95% confidence intervals for sensitivity and specificity. Regression models and statistical tests were performed using SAS[®] software version 8.02 (The SAS System for Windows, SAS Institute, Cary, North Carolina). All statistical tests were conducted at an alpha level of 0.05.

2.4 RESULTS

2.4.1 Search results

The bibliographic database search yielded 516 titles, of which thirty were retrieved (Figure 2.1). A review of reference lists from the primary articles yielded another twelve studies, bringing the total number of studies retrieved and assessed using the inclusion/exclusion criteria to forty-two. Thirty studies were rejected and the reasons for their exclusion are detailed in Appendix IB and outlined in Figure 2.1.

Twelve studies met the inclusion/exclusion criteria. Of these, one study (Jones et al⁵⁸) presented data on caregivers' volunteered history of fever and their impression of body temperature at the time of the interview; only data for the latter were abstracted. In another study (Einterz et al⁴³), only data for children aged 0 to 5 were abstracted and data for children aged 5 to 15 were excluded since several children in this age range (59/176) self-assessed their fever status.

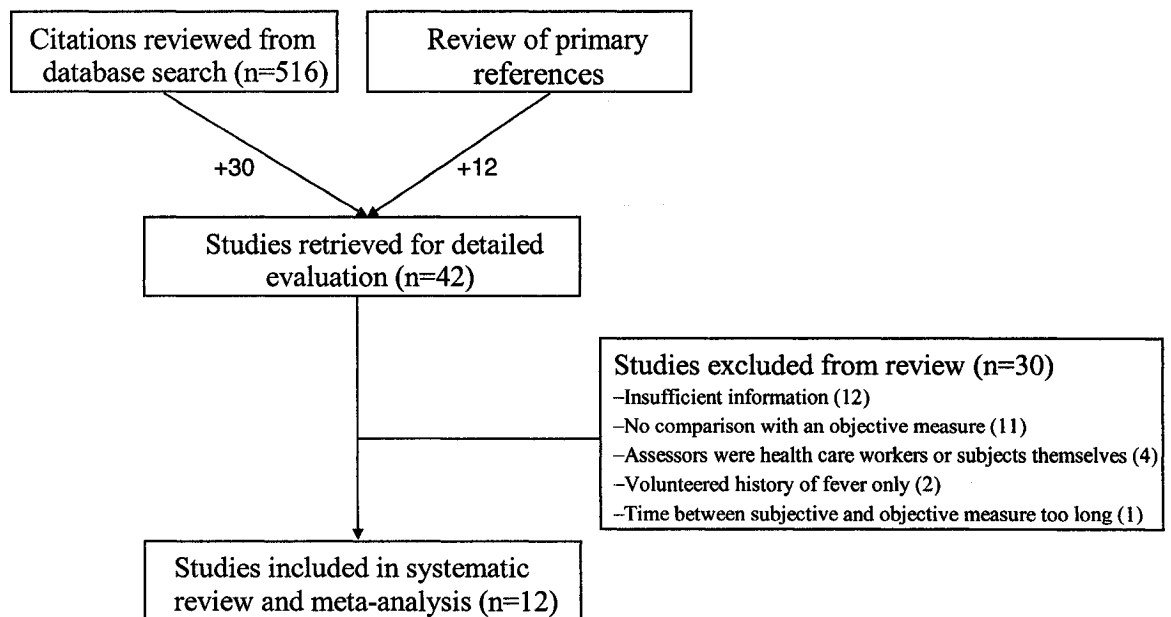


Figure 2.1 Flow diagram of search strategy results

2.4.2 Study characteristics

The characteristics of included studies are presented in Table 2.2. Five studies were conducted in sub-Saharan Africa, four in the United States and one in India, Saudi Arabia and Brazil. The median sample size was 320 and studies ranged in size from 100 to 1,907 children, for a total of 6,606 caregiver-child pairs. Nearly all children and their caregivers were recruited in outpatient clinics or emergency departments; only one study recruited subjects in their homes. The proportion of children with fever as determined by a thermometer exhibited substantial variability across studies but was high overall, with most studies (8/12) recording fever prevalence in excess of 30%. The children were predominantly of pre-school age. Five studies enrolled only children less than five years of age and of those that included older children, the greater part were younger than five years of age. Caregivers were primarily mothers; eight studies enrolled only mothers and the remainder included a small proportion of fathers or other primary caregivers.

The inclusion/exclusion criteria for children and caregivers varied across studies. In one study (Graneto et al⁵⁹), researchers explicitly excluded children who had recently received an antipyretic or other medication from the study, while others included all children regardless of their recent treatment history (Nwanyanwu et al³⁹, Hooker et al⁶⁰). Similarly, some studies excluded caregivers who had used a thermometer prior to their clinic visit (Ernst et al⁶¹, Graneto et al⁵⁹, Banco et al⁶²), while others reported including them (Singhi et al³⁸, Hooker et al⁶⁰, Al-Alamaie et al⁶³). The percentage of caregivers reporting prior thermometer use was quite high in two of these studies (Singhi et al, 31%; Graneto et al, 56%), and low in the two other studies (Al-Almaie et al, 9%; Alves et al, 5%).

Table 2.2 Characteristics of included studies

Study	Country	Setting	Age of Children	No. of children	Site/ Dwell time	Fever definition	Fever prev.
Alves - 2002	Brazil	Clinic	2 mos – 13 years	169	Axilla/N.R.	>38.0°C	81.1%
Al-Almaie - 1999	Saudi Arabia	Clinic	≤ 12 yrs*	1,907	Rectum/N.R. Mouth/N.R.	≥38.3°C ≥37.8°C	20.1%
Whybrew - 1998	Zambia	Clinic	1 month – 16 yrs†	862	Axilla/3min	≥37.8°C	27.4%
Einterz – 1997	Cameroon	Clinic	< 5 yrs	494	Axilla/2min	≥37.5°C	34.2%
Verhoef - 1997	Kenya	Community	2.5 to 37 mos	317	Axilla/N.R.	≥37.5°C	6.0%
Nwanyanwu – 1997	Malawi	Clinic	< 5 yrs	1,120	Rectum/N.R.	≥38.0°C	36.6%
Hooker -1996	United States	Clinic	< 5 yrs	180	Rectum/N.R.	≥38.0°C	55.0%
Graneto - 1996	United States	Clinic	< 10 yrs	322	Rectum/N.R. Tympanic/N.R.	>38.0°C	38.5%
Jones - 1993	The Gambia	Clinic	< 5 yrs	573	Rectum/N.R.	≥38.3°C	30.5%
Singhi - 1990	India	Clinic	3 mos to 12 yrs§	301	Axilla/N.R. Mouth/N.R.	≥37.4°C	38.9%
Ernst - 1985	United States	Clinic	1 mos – 18 yrs	100	Rectum/N.R. Mouth/N.R.	≥38.3°C ≥37.7°C	40.0%
Banco - 1984	United States	Clinic	< 15 yrs	261	Rectum/N.R. Mouth/N.R.	≥38.3°C ≥37.8°C	17.6%

*- median age of 36 months; †- median age of 2 years; § - less than one-third older than five years; N.R. – not reported

2.4.3 Study quality

The results of the evaluation of study quality are presented in Table 2.3. All studies were conducted prospectively and all children who were assessed by a caregiver had their temperature measured with a thermometer. The one possible exception to note is the study by Alves et al⁶⁴ in which children believed to be feverish were purposively sampled, although it is not clear whether this was according to the mother or a health worker and this could have created a bias. The majority of studies performed caregiver assessments blinded to the reference standard, but only one study (Banco et al⁶²) performed both assessments blinded. Deficiencies in reporting of reference standard methodology and results were prevalent. Only two studies (Whybrew et al⁶⁵; Einterz et al⁴³) provided adequate details on the reference standard techniques, with most studies

(10/12) failing to indicate the length of time the thermometer was held in place. Seven studies employed the accepted gold standard of rectal measurements, alone or in combination with another site, for the reference standard while the remaining studies took measurements at the axilla, mouth or a combination thereof. Only two authors (Nwanyanwu et al³⁹; Whybrew et al⁶⁵) provided the reason for not taking rectal temperatures, stating that rectal measurements was not permitted in their facilities. Most studies described how caregivers were asked about their child's fever status, but only seven indicated how caregivers actually performed this assessment in terms of site of subjective assessment. Errors were found in three of the twelve included studies (25%): one reported sensitivity and specificity values and the false negative rate incorrectly (Al-Almaie et al⁶³), one the false positive rate (Nwanyanwu et al³⁹), and one other the false negative rate (Singhi et al³⁸). Only one study (Alves et al⁶⁴) included 95% confidence intervals.

Table 2.3 Results of study quality evaluation

Study Author	Quality Aspect			
	Selection of subjects	Independence of measurements	Reference details	Test details
Alves - 2002	Non-consecutive	Caregiver blind	Insufficient	Sufficient
Al-Almaie - 1999	Not reported	Caregiver blind	Insufficient	Sufficient
Whybrew - 1998	Not reported	No blinding	Sufficient	Sufficient
Einterz - 1997	Non-consecutive	Caregiver blind	Sufficient	Sufficient
Verhoef - 1997	Random	No blinding	Insufficient	Sufficient
Nwanyanwu - 1997	Random	Reference blind	Insufficient	Sufficient
Hooker - 1996	Non-consecutive	Caregiver blind	Insufficient	Sufficient
Graneto - 1996	Consecutive	Caregiver blind	Insufficient	Insufficient
Jones - 1993	Non-consecutive	Caregiver blind	Insufficient	Insufficient
Singhi - 1990	Not reported	Caregiver blind	Insufficient	Sufficient
Ernst - 1985	Consecutive	Caregiver blind	Insufficient	Insufficient
Banco - 1984	Consecutive	Both blind	Insufficient	Sufficient

2.4.4 Diagnostic accuracy

Main Analysis: Table 2.4 presents detailed diagnostic data for each study.

Sensitivity values ranged from 0.74 to 0.97 and specificity values from 0.19 to 0.91.

Values for DOR ranged from 8.59 to 62.10, with the majority being less than 25. Figure 2.2 presents graphically the sensitivity and specificity values and 95% confidence intervals for each study arranged in descending order by sensitivity. The inverse relationship between sensitivity and specificity and the greater variation across studies in specificity values are evident.

Most fevers missed by caregivers were low-grade fevers. Singhi et al³⁸ found that 8 of the 13 children with fever undetected by their caregivers had low-grade fever and in the study by Nwanyanwu et al³⁹ most of the 11 false negatives were cases of low-grade. Similarly, Ernst et al⁶¹ reported that no child with rectal temperature greater than 38.7°C was declared non-febrile by their caregiver.

Table 2.4 Summary of diagnostic information for included studies

Study Author	TP	FN	FP	TN	Sensitivity (95% CI)	Specificity (95% CI)	DOR (95% CI)
Alves - 2002	104	33	3	29	0.76 (0.68-0.83)	0.91 (0.74-0.98)	30.5 (25.5-37.4)
Al-Almaie - 1999	320	64	253	1270	0.83 (0.79-0.87)	0.83 (0.81-0.85)	25.1 (23.2-27.0)
Whybrew - 1998	221	15	353	273	0.94 (0.90-0.96)	0.44 (0.40-0.48)	11.3 (9.3-13.5)
Einterz - 1997	156	13	182	143	0.92 (0.87-0.96)	0.44 (0.39-0.50)	9.4 (6.9-12.0)
Verhoef - 1997	17	2	60	238	0.89 (0.66-0.98)	0.80 (0.75-0.84)	33.7 (28.5-38.9)
Nwanyanwu - 1997	399	11	574	136	0.97 (0.95-0.99)	0.19 (0.16-0.22)	8.6 (7.0-10.2)
Hooker - 1996	81	18	19	62	0.82 (0.73-0.89)	0.77 (0.66-0.85)	14.7 (9.5-19.9)
Graneto - 1996	104	20	48	150	0.84 (0.76-0.90)	0.76 (0.69-0.81)	16.3 (12.2-20.3)
Jones - 1993	155	20	165	233	0.89 (0.83-0.93)	0.59 (0.54-0.63)	10.9 (8.4-13.5)
Singhi - 1990	104	13	21	163	0.89 (0.81-0.94)	0.89 (0.83-0.93)	62.1 (56.6-67.6)
Ernst - 1985	36	4	16	44	0.90 (0.75-0.97)	0.73 (0.60-0.84)	24.8 (16.3-33.2)
Banco - 1984	34	12	31	184	0.74 (0.59-0.85)	0.86 (0.80-0.90)	16.8 (12.3-21.4)

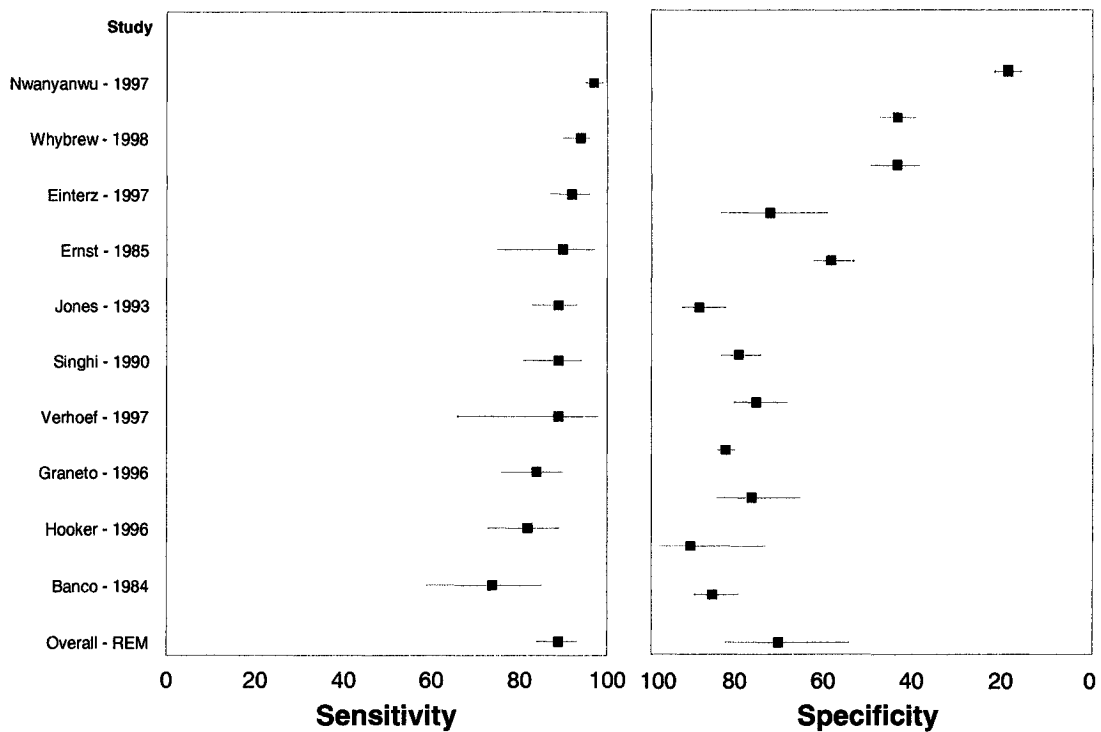


Figure 2.2 Plot of sensitivity and specificity with 95% confidence intervals and the overall pooled estimate using the random effects model (REM)

SROC Curve estimation: A strong positive correlation (Spearman's correlation coefficient = 0.78; $p = 0.003$) between TPR and FPR was found, indicating that a SROC curve should be constructed. Four studies (Jones et al⁵⁸, Whybrew et al⁶⁵, Einterz et al⁴³, and Nwanyanwu et al³⁹) had unusually high FPRs (range 0.41 to 0.81). Consequently, two regression analyses were performed – the first including all studies and the second excluding these four outlying studies.

Figure 2.3 shows the plot of D versus S for both analyses. The line will be approximately horizontal if diagnostic accuracy is constant across the threshold used by caregivers to detect fever. When all studies are included, caregiver accuracy was dependent on the threshold and lower among the four outlying studies where sensitivity is favoured over specificity (high S values and comparatively low D values). This threshold

variation explained a large proportion ($R^2=0.48$) of interstudy differences. When these studies were removed, caregiver diagnostic accuracy was independent of the threshold.

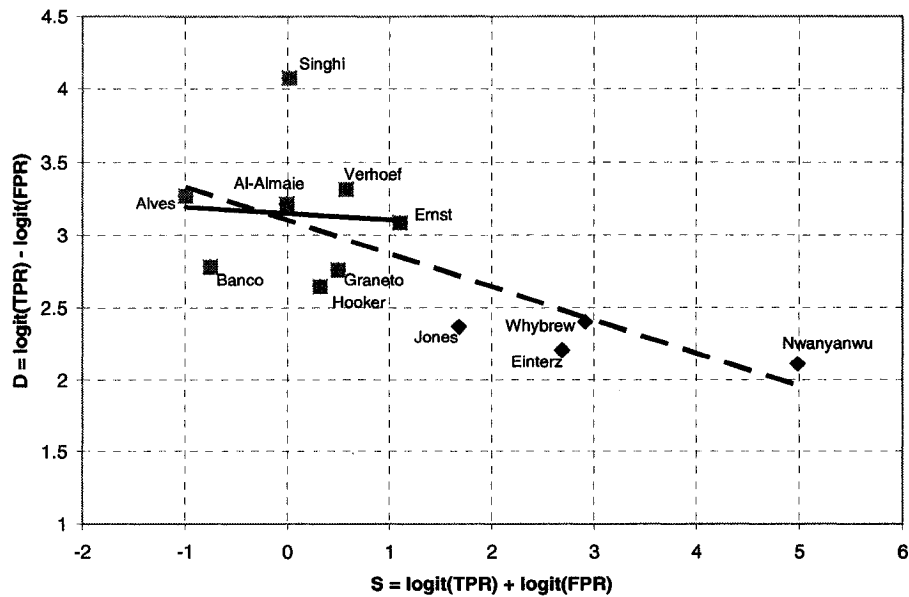


Figure 2.3 Plot of D versus S with the unweighted ordinary least squares regression line for the all studies (dashed line) and outliers removed (solid line) analyses

Parameter and standard error estimates for the slope and intercept for the two analyses (all studies and outliers removed) are shown in Table 2.5. As expected, the intercept is slightly higher when outliers are removed, although this difference was not statistically significant. The regression coefficient for S was statistically significant from zero in the analysis including all studies but near zero when outliers were excluded. Since in both cases the absolute value of S was less than 0.5, suggesting probable symmetry, SDORs were calculated.

Table 2.5 Regression model parameters (weighted and unweighted)

Method	Intercept (α)	Slope (β)	<i>p</i> value	SDOR (95%CI)	A (95% CI)
<i>All studies:</i>					
Unweighted	3.10	-0.23	0.01	22.26 (16.69-29.68)	0.83 (0.80-0.85)
Weighted	3.10	-0.25	<0.01	22.18 (17.35-28.36)	0.82 (0.81-0.84)
<i>Outliers removed:</i>					
Unweighted	3.15	-0.04	0.87	23.30 (16.54-32.83)	0.83 (0.80-0.85)
Weighted	3.15	-0.12	0.74	23.28 (17.40-31.15)	0.83 (0.81-0.85)

Figure 2.4 presents, for both analyses, the scatterplot of TPR versus FPR along with their respective 95% CIs in ellipsoid form with the unweighted (upper) and weighted (lower) SROC curves superimposed. In the all studies analysis, the bulk of studies are clustered together in the upper left hand corner bounded by 70% sensitivity and 70% specificity. The four studies with high false positive rates appear clearly as outliers where caregivers strongly favour sensitivity over specificity. When these studies are excluded from the analysis, the SROC curve remains similar, shifting only slightly towards the upper left hand corner.

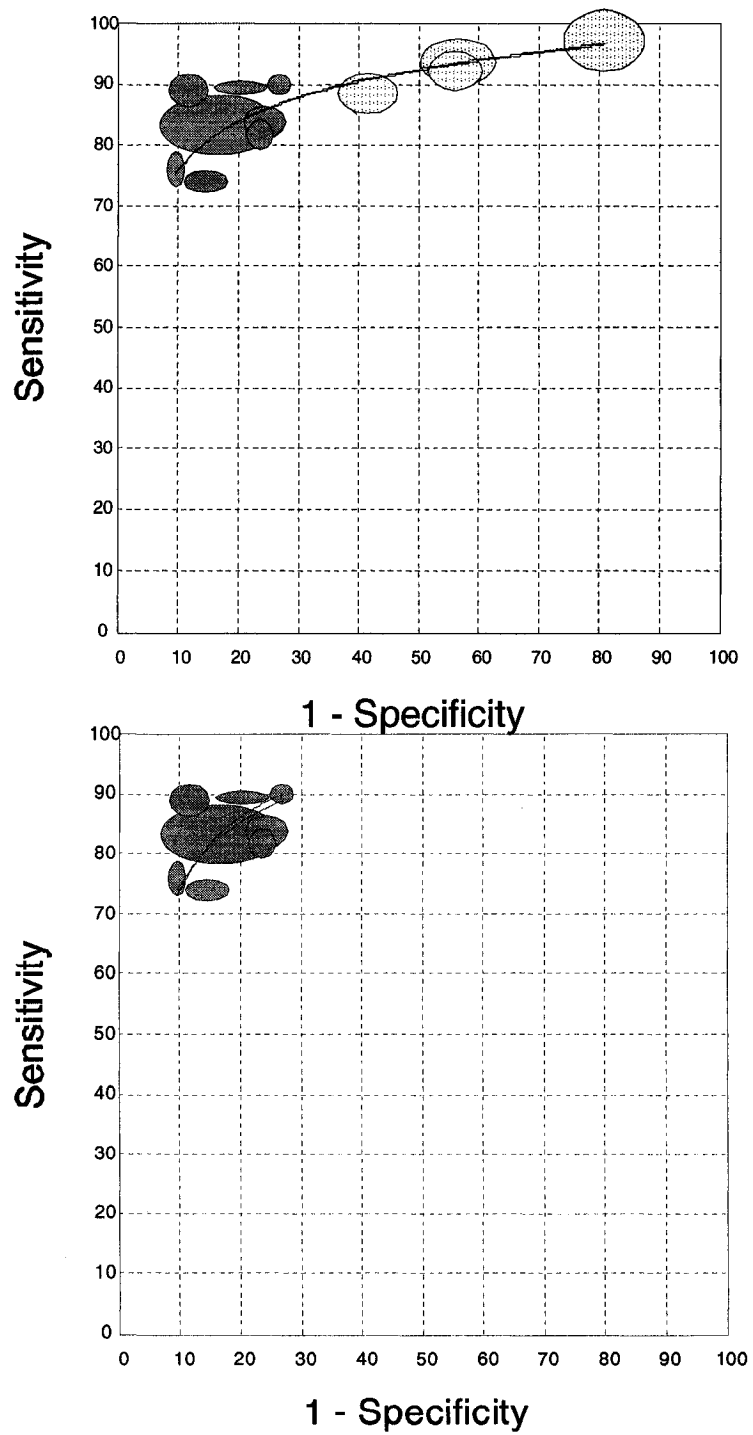


Figure 2.4 Plot of sensitivity against 1-specificity with 95% confidence intervals and the unweighted (upper) and weighted (lower) summary receiver operating characteristic (SROC) curves for the all studies and outliers removed analyses

Estimates of overall test accuracy: The joint value of sensitivity and specificity derived from the unweighted SROC curve was 0.83 for both analyses (Table 2.5). An alternate estimate of overall accuracy used the pooled sensitivity and specificity values to identify corresponding values for the remaining parameter using the unweighted SROC curve. The sensitivity estimate of 0.88 corresponds to specificity values between 65-70% in the all studies analysis and between 75 and 80% in the outliers removed analysis. For the pooled estimate of specificity of 0.71, the corresponding sensitivity estimate is around 85-90% when all studies are considered and approximately 90% after removal of the four outliers.

Subgroup Analysis: Potential sources of heterogeneity examined were: test details (sufficient, insufficient); reference test site (rectal, non-rectal); study region according to magnitude of malaria risk (Sub-Saharan Africa vs. Other); and fever prevalence. Independence of measurements, verification of test results, adequacy of reference details and study setting (clinic vs. community) had less than three studies in each subgroup and could not be examined. No variable attained statistical significance when all studies were included (Table 2.6). When outliers were excluded however, the diagnostic accuracy of caregivers was greater when their child's temperature was measured in the axilla or a combination of sites than when rectal temperatures were measured. The influence of test details reporting and study region could not be explored in the analysis with outliers removed due to the reduced number of studies available.

Table 2.6 Regression model parameters for the subgroup analyses

Characteristic	Method	Intercept (α)	Slope (β)	Parameter (θ)	<i>p</i> value
<i>All studies*:</i>					
Test details	Unweighted	2.95	-0.22	0.18	0.56
	Weighted	2.83	-0.22	0.31	0.25
Reference site	Unweighted	3.29	-0.23	-0.33	0.20
	Weighted	3.28	-0.26	-0.23	0.37
Study region	Unweighted	3.12	-0.20	-0.13	0.76
	Weighted	3.14	-0.13	-0.45	0.29
Fever prevalence	Unweighted	3.17	-0.23	-0.19	0.80
	Weighted	3.17	-0.24	-0.25	0.78
<i>Outliers removed:</i>					
Reference site	Unweighted	3.57	-0.10	-0.69	0.05
	Weighted	3.76	-0.02	-0.72	0.08
Fever prevalence	Unweighted	3.19	-0.06	-0.10	0.91
	Weighted	3.23	-0.11	-0.29	0.79

*Coding: Test details: Sufficient = 1, Insufficient = 0; Reference Site: Rectal = 1, Non-rectal = 0; Study Region: Sub-Saharan Africa = 1, Other = 0.

Additional factors potentially influencing diagnostic accuracy: Several patient and test factors, which could not be examined formally via the SROC curve, were examined within some of the primary studies. Child age and site of subjective assessment were found to be significant in several studies. Child sex, caregiver education level and prior use of a thermometer were not found to be associated with caregiver accuracy^{38;60}.

Child's age: Caregiver accuracy was found to be improved among those assessing younger children, especially those under two years of age. Banco et al⁶² reported greater sensitivity of caregivers to detect fever among children less than two years of age (90% sensitivity vs. 74% for all ages). Ernst et al⁶¹ reported similar results; caregivers of children less than two years old were 93% sensitive and 84% specific in detecting fever compared to the group of 40 caregivers with children between two and eighteen years who were 83% sensitive and 61% specific. Singhi et al³⁸ reported significantly greater

sensitivity but similar specificity among mothers of children less than five years compared to those with children five years and older (94% sensitivity and 76% sensitivity, respectively). Einterz et al⁴³ found caregiver assessment in children less than one year of age was similar to that in children aged one to five years for both sensitivity and specificity.

Site of subjective assessment: Singhi et al³⁸ found caregivers who assessed their child's temperature by touching their child's forehead or face were less sensitive but more specific compared to those palpating the neck or abdomen (80.3% sensitivity and 92% specificity compared to 100% sensitivity and 62.0% specificity, respectively). Caregivers palpating more than one site performed the best, with a sensitivity of 100% and a specificity of 92%. Banco et al⁶² reported that palpation of the trunk and abdomen appeared more accurate than other sites (71.4% positive predictive value for trunk and abdomen compared to 52.3% for all sites combined), but the improvement was not statistically significant due to the small numbers. However, Hooker et al⁶⁰ who compared forehead, forehead plus another body part, trunk or other, and Graneto et al⁵⁹ who examined the correlation between anatomical site palpated (forehead, trunk, arms/legs, abdomen, non-specified) and parental ability to recognize fever found no significant difference. Unfortunately, these studies did not present sensitivity or specificity values according to site, making it difficult to determine whether the results are consistent with the other studies.

2.5 DISCUSSION

This systematic review identified, appraised and combined twelve studies to summarize the accuracy of caregivers' subjective assessment of fever in their children. This discussion will comment on the quality of the primary studies, the main findings, potential factor influencing diagnostic accuracy, the application of results in clinical and community settings and identify some limitations.

2.5.1 Quality of primary studies

Consistent with several reviews and meta-analyses of studies on diagnostic accuracy, the quality of primary studies was generally quite poor⁵³. Positive attributes of the included studies were their reasonable size, prospective data collection, and general avoidance of verification bias. Considerable deficiencies, however, were noted both in the design and reporting of study methodology. Many studies failed to blind both caregivers and health workers, to sample consecutive or random patients, and to provide sufficient details regarding how temperature was measured and how caregivers assessed their children. Based on available information, there were widespread inadequacies in temperature measurement against which caregiver assessments were judged. Few studies measured temperature with the gold standard of rectal measurement and when information was provided, few adhered to the recommended standards for thermometer use, such as calibrating instruments and using correct dwell times. In particular, the use of axillary measurements with insufficient dwell times is likely to have introduced significant error into the measurement of fever. Given that less than one-fifth of axillary temperatures reach their maximum by five minutes, the short dwell times of two to three minutes could have underestimated the number of children with fever and caregiver

specificity. The use of an imperfect reference standard has been shown to introduce bias by overestimating estimates of diagnostic accuracy, since the reference standard and test are generally positively correlated^{54;66}. There is no generally accepted correction for this bias⁵⁴.

Poor reporting of results was also commonplace, particularly with respect to failure to report confidence intervals, which has been noted in other reviews of diagnostic studies⁶⁷. The omission of such simple and easily obtainable estimates of precision for estimates of diagnostic accuracy is important to avoid, since they are often based on rather small sample sizes⁶⁷. Errors in reporting of diagnostic accuracy were found in 25% (3/12) of the included studies, which was better than expected based on a review of diagnostic studies in which 48% reported incorrectly or miscalculated false positive rates⁶⁸. The recent release of reporting guidelines for diagnostic accuracy studies (the Standards for Reporting of Diagnostic Accuracy (STARD) initiative) should be widely publicized and adopted by journals to help improve the quality of reporting of design and results for diagnostic studies⁶⁹.

The overall impact of these design and reporting shortcomings on the estimated accuracy of caregivers is difficult to determine since they operate concurrently and potentially in opposing directions. A recent review of the impact of design flaws on diagnostic accuracy provides some guidance regarding potentially important flaws and their direction of effect⁵³. Inadequate blinding and poor reporting of details regarding test techniques overestimate the DOR while poor reporting of reference standard techniques tend to underestimate the DOR⁵³. Data collection and subject selection methodology, however, appear to have little impact⁵³. The small number of studies in the current

review precluded an exploration of how these design flaws act independently on the estimates of caregivers ability to identify fever and results of the subgroup analysis were inconclusive.

2.5.2 Interpretation of main findings

Despite the shortcomings associated with the primary studies, the synthesis of data regarding caregiver assessments of fever provides new insight into the ability of caregivers to identify fever and some of the factors that influence this ability. Overall, these studies support the value of caregivers' subjective detection of fever. Caregivers are adept at detecting the presence of fever and rarely miss fever in their children as evidenced by sensitivity values exceeding 80% in nearly all studies. Further, most fevers missed by caregivers are low-grade fevers and should such fevers progress, caregivers would most likely be able to identify them. The ability of caregivers to rule out fever correctly is less straightforward as estimates of specificity exhibited greater variability across studies. However, in the majority of studies (8/12), caregivers performed well, identifying correctly more than 70% of children without fever as being afebrile. The greater variation in specificity estimates was due largely to four studies reporting unusually high false positive rates over 40%.

Estimates of sensitivity and specificity are dependent upon each other and must be viewed jointly to understand the ability of caregivers to correctly identify fever. The overall accuracy of caregivers was measured by the SROC curve. Observation of the location of the SROC curve and the distribution of studies around it helps better understand caregiver performance. The close proximity of the curve to the upper left hand corner and the value of joint maximum sensitivity and specificity of 83% obtained

from this curve indicate a strong overall diagnostic performance of caregivers. This accuracy estimate represents an underestimate of the true value since the continuity correction exerts a downward bias^{46;56}. All included studies were located near the SROC curve indicating it is an appropriate summary of the included reports. This proximity to the curve also reveals that a large part of between-study variation is due to, or mediated by, changes in the threshold used by caregivers. Further, the clustering of studies in the upper left hand corner indicates that in most cases, caregivers operate using a threshold that optimizes both sensitivity and specificity.

The ability of caregivers to discriminate between children with fever and children without fever was lower in the four studies distributed at the tail end of the SROC curve where caregivers used a threshold that favoured sensitivity over specificity (DOR range 8.6-11.3 compared to DOR range 14.7-62.1). When these four studies were excluded, a summary DOR was valid, which demonstrated that across the range of sensitivity and specificity estimates, children with fever were more than twenty times as likely to have their caregiver report they were febrile than were children without fever.

The four studies reporting unusually high false positive values shared some common features that shed light on possible reasons behind the observed over-diagnosis of fever and lower discriminatory ability of caregivers in these studies. All four studies were conducted in outpatient clinics in malaria endemic regions of Africa. Fever associated with malaria is often cyclical and body temperature can fluctuate significantly in a matter of hours³⁵. In addition, fever in these areas can often proceed to serious illness within a short time frame. Caregivers may be aware of the tendency for fever to fluctuate and the high risks of overlooking fever and be inclined to report fever based on earlier

examinations. As well, many children identified as feverish by their caregivers may have already been administered an antimalarial such as chloroquine, which acts quickly to reduce fever, and these children were not excluded³⁹. The term fever or 'hot body' as it is often termed in local African languages, may not translate directly into the predominately western concept of fever as a rise in body temperature⁷⁰. Instead, the term may have an alternate meaning that encompasses a broader range of symptoms, which would contribute to the low specificity⁷⁰. Caregivers may also have been concerned that delivery of treatment or care would be withheld if they denied the presence of fever³⁹. The tendency to misreport antimalarial use for similar reasons has been reported in similar populations suggesting that caregivers may knowingly misreport symptoms or behaviours based on their perceptions of health care provider expectations and likely reactions^{71;72}. In a setting where caregivers likely travelled long distances and waited in long line ups to receive care, this is a valid response. These factors help explain the observation of a threshold altered to favour relatively high sensitivity at the expense of specificity reported in these studies.

The two analyses, the one including all studies and the one excluding four studies with high FPR values, produced largely similar results in the summary of caregiver accuracy. High TPR or FPR values can heavily influence the shape of the SROC curve, but generally only when they are characterized by simultaneously high TPR and low FPR or vice versa⁵¹. The high values for FPR observed in this review corresponded to high TPR values. These studies generally followed the SROC curve and do not qualify as true outliers. Discarding those studies falling outside a clinically relevant range is expected to positively bias the SROC curve, as was the case here, but the effect was minimal⁴⁶.

Inclusion of these studies did not meaningfully change estimates of the SDOR and overall accuracy (A), which are largely unaffected by the slope of the regression line. The symmetry of the SROC curve including all studies and the relatively small impact of excluding the studies on measures of accuracy, show the concordance of these studies and their inclusion provides a more complete picture of caregivers' assessments. Relying on the statistical significance of the slope of the regression line may be too sensitive at detecting deviations from a symmetrical SROC curve.

2.5.3 Factors influencing diagnostic accuracy

Threshold variation explains a fair portion of, but not all, between-study variation in diagnostic accuracy and subgroup analysis can help explore sources of residual variation. Unfortunately, the subgroup analysis in this review was limited and confined to univariate analysis due to the small number of studies and the poor reporting on many of the characteristics of interest. Study quality factors including reporting of test details and reference site were not significant influences after adjusting for threshold variation when all studies were included. When outliers were removed, greater diagnostic accuracy was associated with caregivers whose assessments were not compared to the gold standard of rectal temperature. However, these results are inconclusive since other factors could not be controlled for. The prevalence of fever was not shown to have an effect although one might be expected; other studies have shown when there is an imperfect reference standard, sensitivity increases and specificity declines leading to a decrease in the DOR with increasing prevalence⁶⁶.

It is likely that the accuracy of caregiver assessment was also influenced by factors that could not be examined in the subgroup analysis, such as study setting, spectrum of

children examined and caregiver characteristics and perceptions regarding the importance of fever as a symptom. Comparisons within several primary studies suggest that caregivers are better able to identify fever in young children, particularly those less than two years of age. Palpation technique was also associated with diagnostic accuracy in some studies but the results are inconclusive and it is unclear if these results can be generalized beyond the specific populations in these studies. The suggestion from several primary studies that caregivers performed equally well across education levels is encouraging for populations where education levels are generally low. In contrast, mothers with lower education level are less competent in using a thermometer than mothers with higher education levels⁴⁰. Using a thermometer correctly requires a more complex set of skills (i.e. numeracy, knowledge of fever cut-off values) than does palpation, which might explain a greater role of education level in this method of fever detection⁴⁰.

2.5.4 Applications for clinical and community settings

The value of a given test depends strongly on the setting and population in which it is used, the purpose for which it is intended, and concomitant factors such as clinical signs, other tests, and the underlying spectrum of disease⁵⁴. Predictive values, which express the probability a child has fever given a caregiver report of fever are of greater relevance for health providers and programs planners⁵⁵. An important determinant of both positive and negative predictive values is the underlying prevalence of fever in the population of children being evaluated⁵⁵. The studies reviewed here indicate that the prevalence of fever among children visiting outpatient clinics and emergency rooms usually lies between 20% and 40%. Based on these estimates and using the overall

accuracy value of 83%, between 55% and 63% of children identified by their caregivers as feverish would have fever (positive predictive value) and between 83% and 95% of those identified as not having fever would be without fever (negative predictive value).

Caregivers perform well compared to a thermometer, as demonstrated here, and also compared to health professionals in similar settings. Two of the primary studies (Whybrew et al⁶⁵ and Nwanyanwu et al³⁹) directly compared the ability of mothers and health workers and found similar or higher sensitivity but somewhat lower specificity. Potential factors contributing to the poor specificity of caregivers in these two studies have been discussed previously and under different conditions, one might expect specificity values near those of health care workers. Two other studies assessing the ability of health professionals to subjectively assess fever also support the assertion that caregivers perform within the range of health care workers. In a Gambian outpatient clinic, health workers were 65% sensitive and 67% specific and physicians were 79% sensitive and 75% specific and in Kenya, health care workers were 89% sensitive and 92% specific and physicians were 91% sensitive and 77% specific^{73;74}. These results are reassuring for health care workers and others who rely on caregiver assessment.

At a community level, the prevalence of fever among children is generally lower than among children visiting a health facility, usually within the range of 5 to 20% in malaria endemic areas^{23;27;28;30}. Predictive values calculated as before indicate that only about 10% of children identified as feverish by their caregiver would actually be febrile in a low prevalence area (5%) and 55% in a higher prevalence area (20%). Caregivers' ability to identify children without fever, however, would approach 100% (95 to 99%). However, these largely clinic-derived estimates likely underestimate the positive

predictive values in a community setting. When prevalence is low, specificity values rather than sensitivity values largely determine positive predictive value estimates. In a clinical setting, one would expect heightened sensitivity and poorer specificity since the caregiver clearly suspects illness and considers it serious enough to seek care. And as noted earlier, over-reporting of illness may be a strategy employed by mothers to ensure they receive attention and care for their child. At the community level where these conditions are generally absent, the specificity of caregiver assessments would be expected to be higher. The one community study included here simulated at least partly conditions in a clinic since interviews were conducted by medical staff who dispensed treatment for fever, possibly prompting attempts by caregivers to obtain free medicines and may explain the moderate specificity values^{30;75}.

2.5.5 Limitations

The potential for publication bias is a limitation common to meta-analyses⁷⁶. The magnitude of publication bias may be greater for studies of diagnostic accuracy than for randomized controlled trials since many diagnostic studies are small and unlikely to be published unless they produce notably high or low estimates of diagnostic accuracy^{50;77}. Numerous studies, particularly in malaria endemic areas, collect information on a child's temperature as well as a caregiver's assessment of fever^{23;34;78}. However, caregiver accuracy is rarely the focus of the study and diagnostic information is not reported. At present, methods to deal with publication bias for diagnostic studies are not established⁴⁹.

The inclusion of studies where some caregivers had previously used a thermometer may have falsely inflated the assessment of caregivers' diagnostic accuracy. However, the impact of this is expected to be minimal. The one study that did assess this found no

difference in the ability of mothers to identify fever accurately³⁸. Several other studies report that caregivers that use thermometers, often use or interpret them incorrectly and it does not tend to improve their ability to either identify or rule out fever appreciably^{40;41;79}.

The studies included in this systematic review assessed caregivers' ability to identify fever at a single point in time. Consequently, it is not possible to comment on the performance of caregivers over time. Important outstanding information would be the proportion of febrile episodes in a given child that are missed, the nature and outcome of these missed fevers, the number of false fevers detected and the actions taken for such fevers (i.e. an indication of over treatment). This information is particularly important for malarious areas where all fevers are to treated presumptively with an antimalarial and concerns are growing about the development of resistance to safe and affordable antimalarials such as chloroquine⁸⁰. Such information could be collected in a prospective study alongside surveillance for malaria and other febrile illnesses.

The generalization of these results beyond a clinical setting should be approached with caution. Only one study recruited children and caregivers in their homes and did so in a manner that at least partly recreated clinical conditions^{30;75}. Caregivers recruited in a clinic consider their child ill enough to require outside care and their children are more likely to have a fever and to have a high fever. As a result, they may be more sensitive to fever and the high levels of sensitivity observed in the primary studies could reflect this. Where caregivers are asked to detect fever in a setting where children are generally healthy and their assessment does not have an anticipated response as in a community setting, one might expect to see a slight shift in threshold in favour of specificity. Further research is needed to explore this.

3.0 SUB-STUDY II: SOCIO-DEMOGRAPHIC AND HEALTH CARE DETERMINANTS OF CAREGIVER RESPONSES TO CHILDHOOD FEVER

3.1 BACKGROUND

In areas of sub-Saharan Africa where malaria is endemic, episodes of acute febrile illness are presumed to be malaria and appropriate management consists of prompt administration of an effective antimalarial, preferably on the first day of fever onset⁴. The timely provision of an effective antimalarial can be achieved by early home treatment with an antimalarial and/or promptly seeking care from an appropriate provider and both approaches are promoted by the WHO Roll Back Malaria (RBM) strategy⁴. The success of these approaches relies in large part on caregivers' responses to fever. This sub-study explores the influence of socio-demographic and health care factors on caregiver fever management practices in their young children using data collected as part of the C-IMCI initiative in Uganda.

3.1.1 Overview of caregiver management practices for childhood fever

Studies regarding the management of uncomplicated childhood fever reveal that caregivers consider fever seriously; only a small proportion of caregivers report doing nothing when their child developed a fever (4% in Ghana, 5% in Mali, 0% in Zambia, 5% in Kenya)⁸¹. Home treatment with a mix of modern and traditional medicines is common across malaria endemic parts of sub-Saharan Africa and is often initiated before or instead of seeking outside care^{78;82;83}. In Zambia, illness narratives regarding 154 under-five children with fever in the last three weeks revealed that 80% were given some form of home treatment as the first response and 27% were treated solely at home⁸². A mix of modern and traditional practices were employed as initial treatment, including sponging

or bathing with tepid water (35%), administration of antipyretics (34%), chloroquine (13%) and herbal preparations for drinking or rubbing (15%)⁸². In southern Mali, a study of 399 children between one and five years of age with fever in the last three months reported that caregivers managed 76% at home⁷⁸. Of these, 82% were given modern drug treatments and 90% of these treatments included chloroquine alone or in association with another drug. A household survey in rural south-central Togo reported that more than 80% of all fevers among 507 children aged less than five years old were treated at home with an antimalarial⁸⁴.

Caregivers initiate home treatment early, often within the first two days of fever onset. In Mali, two-thirds started treatment within 24 hours, 21% within two to three days and the rest after three days⁷⁸. A study of 532 mothers of children aged six months to ten years in coastal Kenya found that 67% of home treatments were initiated on the same day or the day after fever onset⁸³. In Togo, 97% of those administering an antimalarial at home began treatment on the first day of symptoms and in Bungoma District, Kenya, 91% began within the first two days^{84;85}. However, compliance with recommended regimens is poor and administration of both under and over dosages is common. In Mali, more than half of children provided antimalarials were given incorrect doses and of these, nearly equal proportions were given more than recommended as given insufficient amounts (36% under and 30% over)⁷⁸.

Caregivers commonly obtain drugs for home treatment from informal drug vendors and shopkeepers or use supplies leftover from a previous illness episode. In coastal Kenya, 73% of rural and 96% of urban shops sold antimalarials and 69% of those providing home treatment obtained drugs from these shops⁸³. In the rural district of

Kibaha, Tanzania, all drug stores and shops in the five study villages dispensed chloroquine and one-half of mothers reported stocking drugs for their children under-five⁸⁶. About one-third of those interviewed had drugs at the time of the survey, of which 73% were chloroquine syrup, 4% chloroquine injections, 5% cotrimoxazole, and 8% analgesics⁸⁶. In eastern Uganda, 75% of those giving initial fever treatments at home obtained drugs from shops and 12% from home stocks⁸⁷. Similarly, mothers from 28 villages in Mali reported obtaining drugs to treat fever from shops (38%), stores linked to a health centre (32%), and pharmacies (19%)⁸⁸.

Home treatment of fever generally precedes but does not preclude seeking care outside the home. The proportion of children with fever taken for care outside the home exhibits considerable variability across and within regions. According to the nationwide 2001 Demographic and Health Survey (DHS) in Uganda, two-thirds of children with reported fever in the previous two weeks were taken to care outside the home⁸⁹. In Zambia, 73% of mothers living in communities near health facilities reported taking their febrile child to a health facility during the course of their fever, but only 16% as the first response⁸². In Togo, only 20% of children were taken to a health facility during their illness⁸⁴. An extensive review of DHS surveys in 22 sub-Saharan African countries reported the proportion of childhood fevers taken for care outside the home ranged from 45% to 78% in Eastern and Southern Africa and from 23% to 85% in Western and Central Africa⁹⁰. Sources of care comprised a mix of public and/or private health facilities; only a small proportion of caregivers visited a traditional healer (range 0.3% in Zimbabwe to 4% in Ghana)⁹⁰. Use of multiple care sources also occurs, particularly as

fever grows severe or is accompanied by other symptoms or fails to respond to initial treatments⁹¹.

The promptness of seeking care outside the home is critical for appropriate fever management in malarious areas. Research suggests that caregivers often delay seeking care outside the home by two or more days of illness onset. The median duration of fever reported by caregivers presenting at facilities in Uganda was 3.6 days, with 60% having been ill for less than three days⁸⁷. In Bungoma District, Kenya, only 51% of 134 children taken to a health facility for fever were seen within two days of illness onset and only 9% were seen on the first day⁸⁵. Similarly, in coastal Kenya, 62% of fevers were brought to a clinic, dispensary or hospital on the second day or later⁸³. In Togo, only 17% of children taken to a health facility were taken on the first day of symptoms⁸⁴.

3.1.2 Factors influencing caregiver fever management practices

Several qualitative and quantitative studies have highlighted various socio-demographic and health care factors that influence caregiver responses to fever and these vary across populations. An analysis of illness narratives from mothers living in the catchment areas of rural and urban health centres in Zambia suggested several factors influencing whether a child with fever would be taken to a health centre including: severity of child's condition and whether it was improving or worsening, availability of drugs at the health centre, perceived quality of provider, distance from health centre, and cost of consultation⁸². Lack of childcare for other children and need for permission of other family members were not cited as concerns in this population. In Mali, caregivers cited inability to pay, knowledge of correct treatment, lack of authorization from father,

long distance to health centre, and having the required drugs at home as reasons for not seeking care for their child's fever⁷⁸.

The Uganda DHS (UDHS) 2000-01 survey found children most likely to be taken to care were those aged 6-23 months, urban residents, children of mothers with higher education levels, and residents of the central region⁸⁹. In Kabale district of southwestern Uganda where malaria transmission is unstable, the severity of illness was the most important factor associated with seeking care promptly at a health facility⁹². Maternal education and household wealth were positively associated, but did not attain statistical significance due to the small sample size. A study of 532 caregivers in coastal Kenya found that children younger than five years were more likely to be taken directly to a clinic; younger children were perceived to be more vulnerable to illness and less able to express their illness and to take tablets⁸³. Maternal age, religious affiliation, education level, and income did not help explain the variation in care seeking practices⁸³. However, in Tanzania caregiver education level was positively associated with promptness of seeking care⁹³. Similarly, in Malawi the education level of the household head was associated with an increase likelihood of seeking care and of treating with an antimalarial in the home^{93;94}. A study of mothers of 123 children under five years of age in Ghana found that fewer children in the community of low socio-economic status were taken to a clinic for care compared to those from a community of higher socio-economic status (27% and 42%, respectively)⁸¹. A review of DHS data across sub-Saharan Africa found a positive relationship between care seeking for fever and wealth at the household and village level and for the amount of experience a community has with malaria; the latter association was observed in Eastern and Southern Africa, but not Western and Central

Africa⁹⁰. In southern Tanzania, children from families of higher socioeconomic status, based on an index of education of household head, income sources, and household assets, were more likely to be taken to a health facility than were those of lower socioeconomic status ($p=0.02$ across quintiles)⁹⁵. Higher socioeconomic status was associated with improved knowledge of danger signs and greater physical access to health services, which may explain some of its influence⁹⁵.

3.1.3 Study rationale

Socio-demographic factors and health care aspects are known to influence child mortality and morbidity by operating through intermediate behaviours that directly influence the risk of morbidity and mortality⁹⁶. However, the pattern of influence of these variables for the management of fever by caregivers is relatively unexplored. The studies reviewed above provide valuable insight into these determinants, but most sample sizes have been small and few have explored associations in a multivariate manner or looked at the relative importance of the varying factors. There is also a need to explore the linkages between home treatment and care seeking behaviours and to assess whether these behaviours cluster. Child morbidity and mortality, including fever and severe malaria, have been shown to cluster at the household and village levels but there is limited information on whether caregiver fever management practices cluster and to what degree^{97;98}. The tendency for behaviours to cluster and the magnitude of this clustering can provide insight into approaches for influencing behaviours. Community and household level interventions to improve home treatment and care seeking behaviours for fever are under development in SSA. Uganda is establishing a national program to improve early home management of fevers that places a greater focus on caregiver and

community resources to help reduce malaria mortality⁹⁹. An improved understanding of the nature and determinants of caregiver fever management practices can help shape strategies to optimize behaviours⁴.

3.2 OBJECTIVES

The primary objective of this sub-study is to explore the socio-demographic and health care factors associated with home treatment and care seeking practices of caregivers for childhood fever. Secondary objectives are to assess the extent to which these caregiver practices cluster within villages and health centre catchment areas and to explore the influence of covariates on the magnitude of clustering.

3.3 METHODS

3.3.1 Study area

Uganda, a former British protectorate, is located in east Africa, bordered by Sudan (north), Kenya (east), Tanzania (south), and Rwanda and the Democratic Republic of Congo (west). Uganda has an estimated population of nearly 24 million and is expanding at a rate of approximately 2.9%¹⁰⁰. Children under five comprise approximately 20% of the population. The official language is English and common secondary languages include Ateso, Luganda, Lugbara, Luo, Runyankole/Rukiga, and Runyoro/Rutoro, reflecting the diverse ethnic composition of Uganda⁸⁹. The climate in Uganda is primarily tropical with two rainy seasons in the south (March to May and September to November) and one in the semi-arid north (April to October)¹⁰⁰. Such an environment is conducive to malaria transmission and the Ministry of Health in Uganda estimates that more than 90% of Ugandans live in areas classified as highly endemic for malaria transmission¹⁰¹. The major form of malaria parasite is *Plasmodium falciparum*, which

accounts for an estimated 95% of malaria cases in children¹⁰¹. The national policy is that all fevers should be treated presumptively as malaria and treated with the first line antimalarial¹⁰¹. A combination of chloroquine and sulfadoxine-pyrimethamine (Fansidar) was recommended by the Ministry of Health in 2000 to replace chloroquine alone as the first line treatment for uncomplicated malaria¹⁰².

Health care services in Uganda are decentralized and managed at the district level through Health Directorates¹⁰³. Each district is further divided into sub-counties, parishes and villages. Health centres are primarily based at the sub-county level and serve several parishes. Cost-sharing was discontinued in government health centres in February 2001⁸⁹. In addition to the formal health care providers, there are numerous non-regulated health care providers, including traditional healers, drug sellers and drug shops. Drug sellers and drug shops are particularly common in urban and semi-urban areas and dispense a wide range of antimicrobials without a prescription, including antimalarials such as chloroquine and Fansidar⁸⁷.

3.3.2 Data source

The data for this analysis were collected between August 2001 and February 2002 as part of an ongoing evaluation of the community component of the IMCI in Uganda. This study is a joint initiative of the Johns Hopkins School of Public Health in the United States and Makerere University in Uganda. The C-IMCI aims to improve key household and community practices that impact on child health outcomes, and to accomplish this within existing service structures. Baseline and monitoring surveys in a total of ten districts are planned to assess the impact of the C-IMCI in Uganda (Figure 3.1). Six of these districts (Kiboga, Kumi, Luwero, Masaka, Masindi and Ntungamo) have been

designated as C-IMCI effectiveness sites and four (Bugiri, Iganga, Mubende and Nebbi) as comparison districts. The current survey is considered a baseline survey since community IMCI activities were not yet implemented during the data collection period¹⁰⁴.

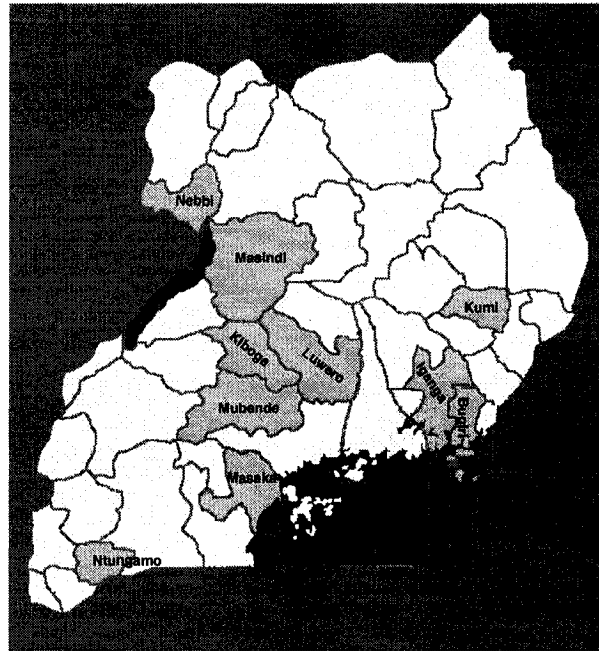


Figure 3.1 Map of Uganda showing districts included in the Community Integrated Management of Childhood Illness (C-IMCI) survey

3.3.3 Sampling methodology

A total of 2,100 households from ten districts were sampled using the modified WHO Expanded Program on Immunization 30 by 7 cluster sampling technique¹⁰⁵.

According to this method, 7 households were randomly selected in 30 village clusters in each of the 10 districts¹⁰⁶. The three-stage stratified cluster sampling procedure used is summarized in Table 3.1.

Table 3.1 Description of sampling units for the Community Integrated Management of Childhood Illness (C-IMCI) survey

Sampling Unit	Selection Process	No. Sampled
District	Purposeful	10
Health Centre	Stratified random	
	Level II	3
	Level III	3
	Level IV	2
Villages	Simple random sampling (from list of villages in 2 or 3 parishes nearest the health centre)	3 or 4 (to total 30 per district)
Caregivers/Children	Systematic random sampling (Consecutive households)	7

Health centre selection: Health centres were the primary sampling unit and eight centres were selected in each district using a stratified approach. In Uganda, health centres are divided into four levels: level I centres have no structure and comprise community outposts; level II centres provide curative and preventative services (i.e. immunization); level III centres also contain a maternity ward; and level IV centres have an operating theatre and inpatient beds¹⁰³. A list of all health centres in each district was obtained and three health centres were randomly selected from level II, three from level III and two from level IV.

Village selection: To provide the sampling frame for selecting villages, the parish containing the health centre and the two or three neighbouring parishes closest to the health centre were identified. All villages within selected parishes were listed and assigned a number. Population data were not available at the village level to allow for proportional-to-size sampling. Consequently, villages were selected using simple random sampling. This procedure was implemented by having a local leader, usually a sub-county chief, select numbers from a table of random numbers.

Caregiver/child selection: In each selected village, study team members presented to the village Chairperson, explained the purpose of the study and asked to be taken to the centre of the village (or if this was unclear or inaccessible, a central meeting place in the village such as a market, school, or large tree). A bottle was placed flat on the ground and spun to indicate the direction to initiate the interviews. The Chairperson was then asked to list the names of all the heads of households in this direction, regardless of whether they had children less than two years of age. Each household head was assigned a number and a table of random numbers was consulted to identify the survey starting point. Consecutive households were then visited until a total of seven eligible households were identified. One caregiver-child pair was interviewed in each eligible household.

3.3.4 Survey respondents

Resident caregivers were interviewed about the health of their youngest child. The survey was restricted to inquiring about children between 0 to 23 months of age to allow the assessment of maternal feeding and antenatal care practices. Mothers were preferred as interview subjects, as they are primarily responsible for the care of young infants and children. When the mother was absent, another adult responsible for the child's care was interviewed.

3.3.5 Study questionnaire

The study questionnaire comprised seven parts, covering socio-demographic characteristics, growth and feeding, home management of fever, care seeking behaviours, disease prevention, knowledge of HIV prevention and health care aspects. The questionnaire was administered in the local language and each interview took approximately 45 minutes to complete. Translation was accomplished by converting the

original English version into local languages and then back translating to ensure the meaning of questions had not changed.

3.3.6 Definition of dependent and independent variables

The subsets of data related to management of fever, socio-demographic characteristics and health care aspects were selected for the current analysis (see Appendix IIA for survey questions). The presence of fever was ascertained by the question “Did this child have fever in the last two weeks?”; only those children whose caregiver answered yes to this question were included. Table 3.2 provides a summary of the dependent and independent variables and indicates their coding.

Table 3.2 Definition and coding of independent and dependent variables

Variables	Definition and coding
<i>Dependent:</i>	
Home treatment with antimalarial	1 = Yes; 0 = No
Outside care sought	1 = Yes; 0 = No
Outside care sought within 48 hours	1 = Yes; 0 = No
<i>Independent:</i>	
Caregiver age	1 = ≤19 years; 2 = 20-29 years; 3 = ≥30 years
Caregiver education level	1 = Less than primary; 2 = Primary or above
Child age	1 = 0-2 months; 2 = 3-11 months; 3 = 12+ months
Child sex	1 = Male; 2 = Female
Household size	1 = <5 members; 2 = ≥ 5 members
Region of residence	1 = Central; 2 = Eastern; 3 = Western; 4 = Northern
Distance to care	1 = <5 km; 2 = ≥5 km
Positive perceived quality of care	1 = Yes; 2 = No
HCW visit in last six months	1 = Yes; 2 = No
Home treatment behaviour*	1 = Nothing; 2 = Antimalarial; 3 = Other

* - predictor variable for the two latter outcome variables only

3.3.6.1 *Dependent variables*

Three main outcome variables were selected based on their relevance to childhood fever management and availability within the dataset. They are defined as follows:

1. Home treatment with antimalarial prior to or instead of seeking outside care.

Home treatment was defined as treatment with an antimalarial before or instead of seeking outside care and determined by self-report in response to the open-ended but probed question “What did you do for the child at home during the illness before seeking advice or treatment outside the home?”. Chloroquine (malariaquine, dawaquine, etc), Fansidar/Metakelfin, camaquine and quinine, administered in any form (tablet, injection, syrup), were considered antimalarials.

2. Outside care sought for fever.

Care seeking was defined as taking the child to care outside the home by the mother herself or another individual and derived using the questions “Did you take the child yourself for treatment outside the home?” and “After how many days or hours of fever did you go to seek care?”. Children were considered as having been taken to care if the caregiver responded yes to the first question OR if there was a time given in response to the second question.

3. Promptness of care seeking.

Care seeking for fever was considered prompt if care was sought within 48 hours of symptom onset and was based on the question “After how many days or hours of fever did you go to seek care?”. Number of days were converted to hours and combined with stated number of hours to give a total time in hours between symptom onset and care seeking. Times were then categorized as less than 48 hours or 48 hours or more. This time period was selected based on evidence that severe complications including death can develop rapidly in children with fever due to malaria, often within two days of symptom onset^{8;26}.

3.3.6.2 *Independent variables*

Socio-demographic and health care variables were selected based on the literature and availability within the dataset. Socio-demographic characteristics included caregiver education level, caregiver age, child age, child sex, household size, and region of residence. Caregiver education level was classified as less than primary or primary and above. Caregiver age was divided into three categories (15 to 19 years, 20 to 29 years, and 30 to 49 years). Child's age was divided into three categories (0-2 months, 3-11 months, 12-23 months) according to the IMCI guidelines for clinical assessment and treatment⁹. A household was defined as a group of people living in the same compound, who prepare and share meals together. Under this definition, a polygamous family may comprise two to three households, with the same man as the head of each of them. The 2000-01 Uganda DHS reported a mean household size of 4.8 members; therefore, household size was divided into less than five and five or more members⁸⁹. Districts were grouped into regions based on Ugandan government classifications: Central (Kiboga, Luwero, Masaka, Mubende); Eastern (Iganga, Bugiri, Kumi), Western (Masindi, Ntungamo); and North (Nebbi)⁸⁹.

Aspects of health care examined included distance to care, perceived quality of care, and extent of community outreach. Distance to external health care source was divided into less than or equal to five kilometres (km) or greater than five km. In Uganda, the Ministry of Health defines reasonable access to a health facility as being within five km of the household¹⁰⁷. Perceived quality of health facility was divided into positive and negative in response to the question "Would you say that the quality of services offered at the health units in your area are good?". To capture information on community outreach activities, respondents were divided into two categories according to

whether they reported receiving a visit from a health care worker (HCW) to their village or household within the last six months.

An additional independent variable was derived to investigate the relationship between care seeking and home treatment behaviours. Home treatment behaviour was divided into three categories: none, antimalarial, and other.

3.3.7 Data entry and quality assurance

At the close of each data collection day, questionnaires were edited and checked for completeness and accuracy, once by the original collector and again by the survey team supervisor. All data were entered into Epi INFO[®] version 6.04 (Centers for Disease Control, Atlanta, Georgia). The author observed the data collection process for the final district (Nebbi, February 2002) and entered the data from this district into Epi INFO[®]. For analysis, data were transferred into SAS[®] version 8.0 (The SAS System for Windows, SAS Institute, Cary, North Carolina). Frequency distributions and cross tabulations were employed to check the data for missing or inconsistent responses, out of range values and other data quality aspects.

3.3.8 Data analysis

This sub-study aims to explore relationships between caregiver fever management practices for fever in their young children and selected socio-demographic and health care aspects and to assess the extent to which such behaviours cluster within villages and within health centre catchment areas. Caregivers living within the same village are likely to be influenced by the behaviours of other caregivers from the same village as well as by shared external influences (i.e. similar source of water, access to transportation) and social and cultural values. These commonalities may result in their responses to fever

being more alike than those of caregivers randomly selected from the population at large. Similarly caregivers residing in the same health centre catchment area, although not perhaps experiencing the same degree of interpersonal contact with other caregivers from the same area, share the same health centre and would receive similar health advice and access to essential medicines, among other common influences. Thus, some within-health centre clustering of caregiver responses to fever in children would be expected, although to a lesser degree than at the village level.

The current data were collected using cluster sampling methodology. This clustering occurred at two main levels – the village and the health centre. This complex sampling design provides both the opportunity to study and the need to address the correlated behaviours of caregivers at both the village and health centre level. Sampling strategies other than simple random sampling bias point estimates and underestimate sample variance, resulting in tests of significance that favour rejection of the null hypothesis (increased Type 1 error)¹⁰⁸. The degree to which the behaviours of caregivers from a given cluster are more similar than that observed in a simple random sample is commonly referred to as the intraclass correlation coefficient (ICC) and directly influences the degree to which sample variance is increased¹⁰⁵. This increased sample variance in complex surveys compared to surveys based on simple random sampling for the same number of respondents is termed the design effect, or inflation factor¹⁰⁵. The design effect is expressed mathematically as $[1 + (n-1)p]$, where n is the number of children recruited per cluster and p is the ICC¹⁰⁵. The ICC is estimated by dividing the between-cluster variation by the sum of the between- and within-cluster variation. When caregivers in each cluster behave independently of each other (i.e. the ICC approaches 0),

the variance will be similar to that calculated assuming simple random sampling. Ordinary logistic regression procedures (OLR) and generalized linear models (GLM), which are used to explore associations for binary outcomes, assume that data are based on simple random sampling and are therefore independent¹⁰⁸. This criterion is not met in the current survey and an approach to account for the correlations within the data is required for a valid analysis. To address this, and to explore the extent of behavioural dependency, a newly developed method of alternating logistic regressions was employed.

3.3.8.1 Alternating logistic regressions

The method of alternating logistic regressions (ALR) was developed in the early 1990s by Carey et al, building on the work of those who created generalized estimating equations (GEE) for the analysis of correlated multivariate data¹⁰⁹. Alternating logistic regressions offer several advantages for the analysis of data with binary outcomes collected using cluster sampling. First, ALR generates estimates for regression parameters that adjust for the study design by accounting for the correlation between responses¹⁰⁹. Second, it provides an interpretable estimate of the extent to which responses are correlated at a given level of aggregation. It is primarily this feature that distinguishes ALR from other approaches; other methods such as GEE simply adjust for the correlational structure of the data, while ALR focuses on providing more efficient estimates of this correlation and expresses it as a pairwise odds ratio (PWOR)¹¹⁰. Third, ALR allows an exploration of the influence of covariates on the magnitude of clustering. Finally, ALR is able to handle both large and small cluster sizes and can incorporate several levels of aggregation^{109;111}.

In technical terms, ALR is a type of marginal regression modelling where the dependency of the outcome variable on covariates is modelled separately from the correlations between outcomes¹⁰⁹. The ALR algorithm stipulates two series of regressions that are performed simultaneously until convergence is reached: 1) estimate the pairwise odds ratio and run a generalized estimating equation logistic regression to obtain regression coefficients for the covariates; and 2) update the pairwise odds ratio using an offset logistic regression that relates paired outcomes within each cluster¹¹¹. This offset regression depends on the values for the regression coefficients and pairwise odds ratios generated in the first series and thus requires iteration of the model¹⁰⁹. The procedures used to obtain the pairwise odds ratios and regression coefficients as applied to the current data for home treatment are discussed in turn below.

a) Estimation of the pairwise odds ratio. Alternating logistic regressions estimates the association between caregiver responses to fever using the pairwise odds ratio, also commonly termed the pairwise cross product ratio¹¹⁰. In this study, the clustering of caregiver behaviours is explored at the level of the village and the health centre catchment area using a nested design, in which villages were nested within health centre catchment areas. The estimation of the PWOR proceeds by taking successive pairs of caregivers from the same health centre or cluster and calculates the odds that members of these groups will have the same outcome¹¹⁰. This can be conceptualized as a two-by-two table with one outcome on the vertical and the other on the horizontal and the number of pairs falling in each of the four cells¹¹¹. Cells A and D represent concordant pairs and cells B and C discordant pairs. The PWOR is then calculated as the cross product (AD/BC) ¹¹¹.

In the ALR model, the PWOR for within village and within health centre catchment areas was calculated using the following equation¹¹²:

$$\text{Log} [\text{PWOR}(Y_{ijk}, Y_{ilm})] = \alpha_0 + \alpha_1 Z_{ijklm}, k \neq m$$

Using the outcome of the home treatment with an antimalarial as an example, Y_{ijk} is the binary outcome variable indicating whether the k^{th} caregiver in the j^{th} village of the i^{th} health centre catchment area administered an antimalarial to her feverish child before or instead of seeking care¹¹². Y_{ilm} represents a randomly selected response from same health centre (i), a potentially a different village (l), and a different caretaker (m). As noted above, k and m are mutually exclusive¹¹². The term Z_{ijklm} is equal to 1 when both caregivers are from the same village ($j = l$) and equal to 0 when the caregivers are from the same health centre but not the same village¹¹². Therefore, the within-village PWOR is equal to $\exp(\alpha_0 + \alpha_1)$ and the within-health centre PWOR is equal to $\exp(\alpha_0)$ ¹¹².

The within-village PWOR represents the odds of home treatment with an antimalarial by a caregiver given that another caregiver randomly selected from the same village has treated with an antimalarial relative to the odds that the latter has not treated their child with an antimalarial¹¹¹. The within-health centre PWOR represents the odds of home treatment by a caregiver from a health centre where another caregiver randomly selected from the same health centre, but not the same village, has treated their child with an antimalarial relative to the odds that the randomly chosen respondent has not treated with an antimalarial. In this way, the within-health centre PWOR essentially indicates the amount of residual clustering in caregiver behaviours after the clustering at the village level has been taken into account¹¹². It should be noted that only those villages and health centres with two or more caregiver-child pairs in the regression analysis contribute to the

estimate of the PWOR¹¹¹. However, all caregiver-child pairs contribute to the estimate of the regression coefficients¹¹¹.

The interpretation of the PWOR is similar to that of an ordinary odds ratio¹¹¹. A PWOR of 1.0 indicates no association between caregiver behaviours at a given level of aggregation and values greater than 1.0 indicate the presence of an association. The degree of clustering increases with the magnitude of the PWOR. The PWOR has several properties which make it attractive for expressing the correlation between binary outcomes. It can take on any value on a real line while other correlation measures, such as the ICC, are influenced by the likelihood of a positive outcome and perform poorly when outcomes are binary^{97;109}. In addition, the PWOR is independent of cluster size and the influence of other covariates on the PWOR can also be easily explored within the ALR framework⁹⁷.

b) Estimation of the regression parameters. The ALR procedure estimates regression coefficients using GEE¹⁰⁹. The equation can be expressed as:

$$\text{Logit} [\text{prob}(Y_{ijk} = 1)] = \beta_0 + \beta_1 X_{1ijk} + \beta_2 X_{2ijk} + \beta_3 X_{3ijk} \dots$$

where, X represents a covariate and β its regression coefficient¹¹². The initial step in the GEE procedure is to model the regression parameters using the GLM approach, which treats responses as independent¹¹³. The GEE then extends the GLM approach to account for the dependent nature of the data via a second step in which the correlations between responses are modelled¹¹³. This is accomplished through the specification of a working correlation matrix, which is usually unknown and must be estimated¹¹³. In the case of ALR, the within-cluster correlation is modelled using the pairwise log odds ratios obtained from the first series of equations¹⁰⁹.

Several advantages of the GEE approach for estimating regression coefficients in the context of ALR are that it reduces to GLM when there is no correlation in the data; does not require that the data be normally distributed, and can accommodate large numbers of clusters of varying sizes¹¹³. In addition, empirical (or robust) estimates of variance are valid even if the working correlation matrix is mis-specified¹¹³. The GEE approach requires a relatively large number of clusters (i.e. more than 20-30) for a valid analysis¹¹³. There are 300 clusters in the current study indicating that this criterion is adequately met. Other assumptions of the GEE approach and thus ALR are that the model is linear and that observations in different clusters are independent¹¹⁴.

3.3.8.2 *Approach to modelling*

Contingency tables were run to examine each outcome variable by the different levels of categorical predictors to identify any zero cells or cell sizes less than five. Following this, three sets of models were performed for each of the dependent variables:

1. Intercept only ALR models to generate the unadjusted within-village and within-health centre pairwise odds ratios;
2. Univariate ALR models to estimate the unadjusted regression coefficients for covariates; and
3. Multivariate ALR models to obtain adjusted parameter values for covariates and within-village and within-health centre pairwise odds ratios adjusted for covariates.

In all three sets, only those caregiver-child pairs with complete information for all covariates were included to facilitate comparisons between unadjusted and adjusted parameter estimates (approximately 79% of eligible data). The ALR algorithm was conducted using the PROC GENMOD procedure in SAS[®] version 8.0. Correlations

among the binary outcomes were modelled by PWORs in which villages were specified as subclusters of the health centre catchment areas. Type 3 (Wald) statistics were obtained for each parameter in the model and used to assess the contribution of each variable to explain the outcome of interest. Empirical (robust) standard errors were used to estimate 95% confidence intervals for parameter estimates. Results were considered statistically significant at an alpha level of 0.05 and all tests were two-tailed. Where applicable, *p* values were reported along with 95% confidence intervals.

3.4 RESULTS

A total of 2,098 caregivers were interviewed, of which 92.7% were the mother of the child, 4.0% the father, 2.2% another adult in household and 1.1% other or unknown. For consistency, the analysis was restricted to maternal caregivers. A total of 55 children were older than 24 months and were excluded from further analysis. Thus, data regarding a total of 1,883 caregivers of children aged 0 to 23 months were available for analysis (Figure 3.2). Of the caregivers who provided information regarding fever, 837 (44.5%) reported that their child had had fever within the last two weeks. Information regarding home treatment and care seeking practices was available for 827 (98.8%).

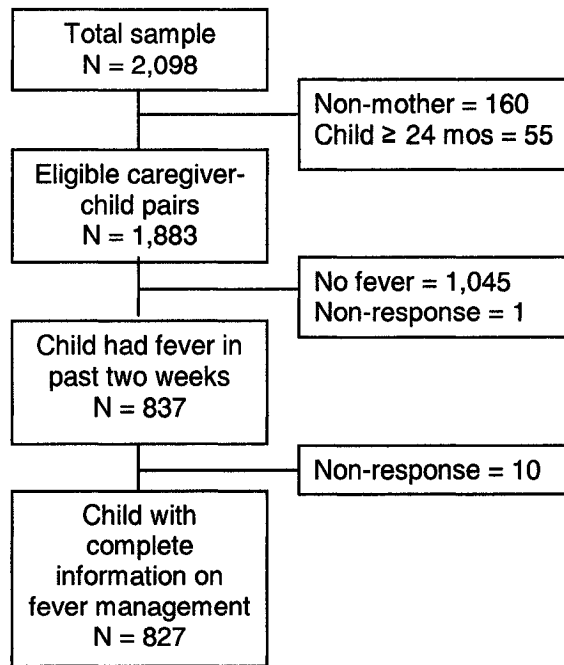


Figure 3.2 Flow diagram of Community Integrated Management of Childhood Illness (CIMCI) survey respondents

3.4.1 Sample characteristics

3.4.1.1 Socio-demographic and health care characteristics

Maternal caregivers ranged in age from 16 to 48 years, with 14.2% aged less than 19 years, 59.8% 20 to 29 years and 26.0% 30 years and older. Most mothers had low education levels: 75.4% reported less than primary and only 24.6% reported completing primary education or more. The majority of mothers (87.6%) were married.

Approximately 5.7% of the children were between 0 and 2 months of age, 44.8% between 3 and 11 months of age, and the remainder one year or older. There were relatively equal proportions of male and female children (51.8 and 48.2, respectively). The median household size was 5.0 and 81.2% reported less than 8 members.

Nearly all caregivers (98.4%) reported seeking care at a modern health facility, with 70.3% using either a government or a non-governmental health facility and 28.1% preferring a private clinic. The remainder (n=13) consulted a traditional healer or other

provider. The median distance from home to care source for those reporting visiting a modern facility was 2.0 km and 78.1 % were within five km. Approximately 62.6% of caregivers perceived the quality of services available as good. Twenty percent of caregivers reported receiving visits by health workers to the village or the household within the last six months.

3.4.1.2 Fever management practices

Table 3.3 presents the self-reported home treatment practices taken by caregivers to manage their child's fever before or instead of seeking care. Of the 645 caregivers (77.0%) who undertook initial home management, 57.1% reported one activity, 36.9% two activities and 5.6% three or more activities. A total of 347 mothers (41.9%) reported giving an antimalarial, of which chloroquine accounted for 91.9%.

Table 3.3 Home care practices for childhood fever before or instead of seeking outside care reported by caregivers

Activity	N (n = 827)	Percent*
Painkillers (Panadol/Paracetamol, Aspirin)	365	44.1
Chloroquine	319	38.6
Nothing	182	22.0
Performed tepid sponging	94	11.4
Unknown drug	76	9.2
Herbal medicine	50	6.0
Other antimalarial (Fansidar, Quinine, etc)	35	4.2
Other	18	2.2

* Caregivers could report more than one activity and as a result, percentages do not sum to 100

Seventy-three percent of children (604/827) were taken for care outside the home at some point during their fever episode. Most (95%) were taken by the caregiver and the remainder by someone other than the caregiver. Table 3.4 provides a summary of the home treatment and care seeking activities and their sequence. Of those children brought

to care, more than one-third (37.4%) had already received an antimalarial. A total of 102 children (12.3%) were not taken to care and did not receive an antimalarial at home. For children taken to care, information on time from fever recognition to presentation at care source was available for 572 cases (94.7%). Of these, 32.5% were brought to care within 24 hours, 24.1% between 24 and 47 hours and 43.4% after 48 hours or more.

Table 3.4 Summary of home care and care seeking practices for childhood fever reported by caregivers

Activity	N (n = 827)	Percent
Home treatment with antimalarial → outside care	226	27.3
Other home treatment → outside care	218	26.4
Outside care	160	19.3
Home treatment with antimalarial	121	14.6
Other home treatment	80	9.7
Nothing	22	2.7

3.4.2 Factors associated with fever management practices

The distribution and number of caregiver-child pairs per village and health centre included in the ALR models for the three outcome variables are displayed in Table 3.5. For the home treatment and care seeking ALR models, 260 villages contributed information regarding a total of 655 individuals (79.2% of children with reported fever during the two weeks preceding the interview). Of these villages, 187 (71.9%) contained two or more caregiver-child pairs and were used in the estimation of the village PWORs. For the promptness of care seeking model, 223 villages were included for a total of 457 individuals (79.9% of all fever cases taken to care). Two or more caregiver-child pairs were available to estimate the PWOR in only 133 (57.1%) villages. For all outcomes,

nearly all health centres had sufficient numbers to contribute to the estimation of the PWORs at that level.

Table 3.5 Number of children per village and health centre cluster

No. of caregiver-child pairs	Home treatment/ Care seeking (n=665)		Promptness of care seeking (n=457)	
	Clusters	%	Clusters	%
<i>Per village</i>				
1	73	28.1	90	40.4
2	62	23.8	65	29.1
3	67	25.8	42	18.8
4	36	13.8	20	9.0
5	19	7.3	5	2.2
6	3	1.2	1	0.4
Total	260	100	223	100
<i>Per health centre</i>				
1	0	0.0	1	1.3
2-5	21	26.3	40	50.0
6-10	41	51.3	34	42.5
>10	18	22.5	5	6.3
Total	80	100	80	100

3.4.2.1 Home treatment with an antimalarial

Several socio-demographic factors including caregiver education level, child age and household size were significantly associated with administering an antimalarial at home before or instead of seeking care in the univariate and multivariate models (Table 3.6). Caregiver age was borderline significant ($p=0.08$) in the univariate model only.

After adjustment for covariates, primary education or greater was associated with a nearly two-fold increase in the likelihood of administering an antimalarial at home (OR: 1.88; 95% CI: 1.25, 2.85). Older children aged 3 to 11 months and 12 months or older were more likely to be treated in the home first with an antimalarial compared to infants less than or equal to two months old. Larger household size (five or more members) was associated with just over a 50% increase in likelihood of maternal administration of an

antimalarial (OR: 1.69; 95% CI: 1.26, 2.27). No health care variables were significantly associated with home treatment with an antimalarial.

Table 3.6 Unadjusted and adjusted odds ratios (OR) and 95% confidence intervals for home treatment with an antimalarial before or instead of seeking care among caregivers reporting a child with fever during the last two weeks

Variable (Reference Category)	Unadjusted†			Adjusted‡		
	OR	95% CI	<i>p</i>	OR	95% CI	<i>p</i>
Caregiver age (≤19 years)			0.08			0.29
20 – 29 years	1.41	0.84, 2.38		0.95	0.54, 1.68	
30 – 49 years	1.64	1.06, 2.55		1.25	0.80, 1.96	
Education level (<Primary)			<0.01			<0.01
≥Primary	1.87	1.27, 2.76		1.88	1.25, 2.85	
Child's age (0-2 months)			<0.01			<0.01
3-11 months	2.97	1.28, 6.86		2.51	1.11, 5.69	
≥ 12 months	1.74	1.57, 4.20		1.47	0.62, 3.52	
Child sex (Male)			0.74			0.85
Female	1.06	0.76, 1.46		1.03	0.73, 1.46	
Household size (<5 members)			<0.01			<0.01
≥ 5 members	1.57	1.18, 2.08		1.69	1.26, 2.27	
Region of residence (Central)			0.33			0.80
Eastern	0.74	0.45, 1.20		0.97	0.57, 1.66	
Western	0.94	0.55, 1.62		1.11	0.65, 1.90	
Northern	1.18	0.74, 1.86		1.23	0.77, 1.94	
Distance to care (≥5 km)			0.66			0.89
<5 km	1.09	0.63, 1.33		0.97	0.66, 1.43	
Positive perceived quality of care (No)			0.92			0.83
Yes	0.98	0.69, 1.39		1.04	0.72, 1.51	
HCW visit in last six months (No)			0.22			0.62
Yes	0.77	0.51, 1.16		0.89	0.55, 1.42	

† Includes only those with complete information on caregiver age, caregiver education level, child age, child sex, household size, region of residence, distance to care, perceived quality of care, and HCW visit in last six month (N=655)

‡ Adjusted for variables listed above

3.4.2.2 Care seeking

Previous home treatment practices, distance to care source, and region of residence were strong independent predictors of seeking care in the multivariate model (Table 3.7). Caregivers who reported giving their child an antimalarial or another form of home treatment were much less likely to seek care compared with caregivers who reported no initial home treatment (OR: 0.18; 95% CI: 0.10, 0.31 and OR: 0.27; 95% CI: 0.16, 0.46, respectively). Living within five kilometres of care source was associated with a two-fold increase in likelihood of seeking care for fever compared to living farther away (OR: 2.01; 95% CI: 1.32, 3.04). Perceived positive quality of health service and reporting of a health care worker visit within the last six months were associated with a slightly increased likelihood of seeking care in the univariate model, but these did not attain statistical significance ($p=0.14$ and $p=0.12$, respectively) and declined in the multivariate model. Caregivers living in the eastern and northern regions of Uganda were significantly more likely to seek care than caregivers in the central region. A similar direction of effect was observed in the western region but the increase was not statistically significant.

Table 3.7 Unadjusted and adjusted odds ratios (OR) and 95% confidence intervals for seeking outside care among caregivers reporting a child with fever during the last two weeks

Variable (Reference Category)	Unadjusted†			Adjusted‡		
	OR	95% CI	<i>p</i>	OR	95% CI	<i>p</i>
Caregiver age (≤19 years)			0.63			0.80
20 – 29 years	0.90	0.54, 1.48		1.09	0.64, 1.85	
30 – 49 years	0.79	0.48, 1.30		0.93	0.55, 1.57	
Education level (<Primary)			0.80			0.24
≥Primary	1.05	0.72, 1.54		1.29	0.85, 1.98	
Child's age (0-2 months)			0.58			0.85
3-11 months	0.84	0.40, 1.76		1.07	0.47, 2.43	
≥ 12 months	1.01	0.45, 2.25		1.19	0.50, 2.84	
Child sex (Male)			0.79			0.91
Female	0.95	0.68, 1.34		0.98	0.68, 1.41	
Household size (<5 members)			0.35			0.94
≥ 5 members	0.86	0.63, 1.17		0.99	0.70, 1.39	
Region of residence (Central)			<0.01			<0.01
Eastern	3.56	2.02, 6.27		3.34	1.73, 6.45	
Western	1.44	0.79, 2.61		1.39	0.78, 2.47	
Northern	1.71	1.15, 2.54		1.88	1.30, 2.71	
Distance to care (≥5 km)			<0.01			<0.01
<5 km	2.06	1.36, 3.11		2.01	1.32, 3.04	
Positive perceived quality of care (No)			0.14			0.35
Yes	1.28	0.93, 1.76		1.18	0.84, 1.66	
HCW visit in last six months (No)			0.12			0.78
Yes	1.41	0.92, 2.17		1.07	0.66, 1.72	
Home treatment (None)			<0.01			<0.01
Antimalarial	0.20	0.12, 0.35		0.18	0.10, 0.31	
Other	0.28	0.16, 0.49		0.27	0.16, 0.46	

† Includes only those with complete information on caregiver age, caregiver education level, child age, child sex, household size, region of residence, distance to care, perceived quality of care, HCW visit in last six month, and home treatment (N=655)

‡ Adjusted for variables listed above

3.4.2.3 *Promptness of care seeking*

Region of residence, distance to care, and home treatment practices were related to promptness of care seeking in both the unadjusted and adjusted models (Table 3.8).

Caregivers providing either antimalarials or some other form of home treatment were considerably less likely to bring their child to care within 48 hours compared with caregivers providing no initial home treatment (OR: 0.55; 95% CI: 0.34, 0.87 and OR: 0.58; 95%CI: 0.37, 0.93, respectively). Caregivers living within five kilometres of a care source were significantly more likely to seek care for their child with fever within 48 hours than their counterparts living five or more kilometres from a care source.

Table 3.8 Unadjusted and adjusted odds ratios (OR) and 95% confidence intervals for promptness of care seeking among caregivers reporting seeking care for a child with fever during the last two weeks

Variable (Reference Category)	Unadjusted†			Adjusted‡		
	OR	95% CI	<i>p</i>	OR	95% CI	<i>p</i>
Caregiver age (≤19 years)			0.11			0.16
20 – 29 years	0.94	0.53, 1.68		1.00	0.54, 1.84	
30 – 49 years	0.65	0.38, 1.11		0.68	0.39, 1.20	
Education level (<Primary)			0.92			0.62
≥Primary	0.98	0.61, 1.58		1.13	0.70, 1.83	
Child's age (0-2 months)			0.49			0.78
3-11 months	0.58	0.24, 1.42		0.73	0.30, 1.78	
≥ 12 months	0.61	0.25, 1.50		0.73	0.30, 1.79	
Child sex (Male)			0.84			0.83
Female	0.96	0.66, 1.41		0.96	0.65, 1.42	
Household size (<5 members)			0.59			0.68
≥ 5 members	0.90	0.61, 1.32		0.92	0.62, 1.37	
Region of residence (Central)			0.01			0.04
Eastern	2.36	1.36, 4.10		2.21	1.18, 4.15	
Western	1.02	0.59, 1.78		1.04	0.60, 1.80	
Northern	1.44	0.88, 2.34		1.61	0.97, 2.69	
Distance to care (≥5 km)			0.01			0.03
<5 km	1.77	1.13, 2.78		1.70	1.06, 2.75	
Positive perceived quality of care (No)			0.93			0.96
Yes	1.02	0.64, 1.63		0.99	0.61, 1.60	
HCW visit in last six months (No)			0.10			0.43
Yes	1.49	0.93, 2.41		1.23	0.74, 2.03	
Home treatment (None)			0.04			0.02
Antimalarial	0.55	0.35, 0.87		0.55	0.34, 0.87	
Other	0.63	0.40, 1.00		0.58	0.37, 0.93	

† Includes only those with complete information on caregiver age, caregiver education level, child age, child sex, household size, region of residence, distance to care, perceived quality of care, HCW visit in last six month, and home treatment (N=457)

‡ Adjusted for variables listed above

3.4.3 Clustering of fever management practices

The unadjusted and adjusted within-village and within-health centre pairwise odds ratios for each outcome are given in Table 3.9 and depicted in Figure 3.3. In the unadjusted model, PWORs were consistently higher at the village level than the health centre level and this difference was greatest for promptness of care seeking. At the village level, the PWORs were statistically significant for all three outcomes. The magnitude of clustering within villages was greatest for promptness of care seeking (PWOR: 2.35; 95% CI: 1.36, 4.05) and lowest for home treatment with an antimalarial (PWOR: 1.21; 95% CI: 0.97, 1.51). The opposite trend was observed for health centres, with the degree of clustering being greatest for home treatment with an antimalarial (PWOR: 1.19; 95% CI: 0.91, 1.55) and close to the null value for care seeking and promptness of care seeking (PWORs: 1.08 and 1.03, respectively). None of the health centre odds ratios reached statistical significance in the unadjusted model.

Adjustment for socio-demographic and health care covariates attenuated both the village and health centre catchment area associations, but in a different pattern. At the village level, clustering was reduced for home treatment and seeking outside care, but not for promptness of care seeking which remained steady at 2.35. At the health centre level, the PWORs at the health centre level was essentially unchanged for home treatment but declined to less than 1.0 for the care seeking and promptness of care seeking and for care seeking, the PWOR was significant (OR: 0.73; 95% CI 0.59, 0.90).

Table 3.9 Unadjusted and adjusted within-village and within-health centre catchment area pairwise odds ratios (PWOR) for home treatment with an antimalarial, care seeking, and promptness of care seeking

	Unadjusted†		Adjusted‡	
	Village PWOR (95%CI)	Health Centre PWOR (95%CI)	Village PWOR (95%CI)	Health Centre PWOR (95%CI)
Home treatment (n=665)	1.46 (1.07, 2.00)	1.21 (0.97, 1.51)	1.19 (0.91, 1.55)	1.17 (0.94, 1.44)
Care seeking (n=665)	1.85 (1.18, 2.88)	1.03 (0.79, 1.34)	1.59 (0.98, 2.57)	0.73 (0.59, 0.90)
Promptness of care (n=457)	2.35 (1.35, 4.07)	1.08 (0.75, 1.56)	2.35 (1.36, 4.05)	0.93 (0.70, 1.23)

† - Unadjusted values obtained from the intercept only model

‡ - Adjusted for maternal age, caregiver education level, child age, child sex, household size, region of residence, distance to care, perceived quality of care, HCW visit in last six month, and home treatment (care seeking and timeliness of care only)

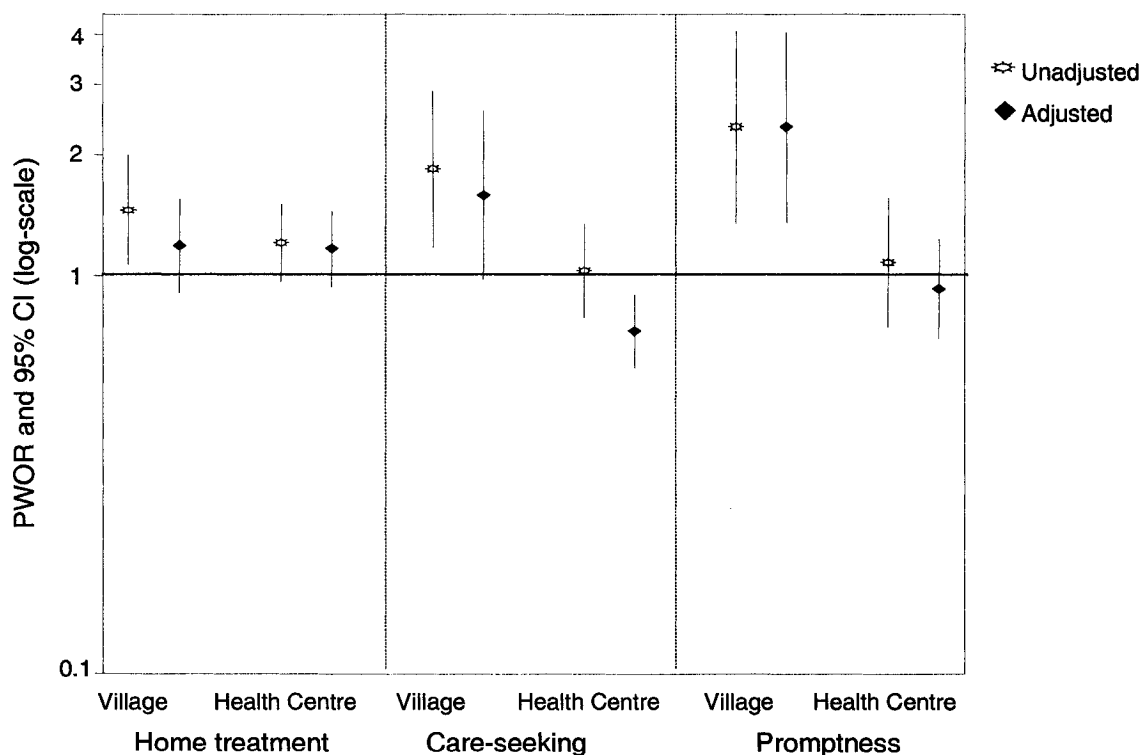


Figure 3.3 Unadjusted and adjusted within-village and within-health centre catchment area pairwise odds ratios (PWORs) for home treatment with an antimalarial, care seeking, and promptness of care seeking

3.5 DISCUSSION

This sub-study employed newly developed techniques of alternating logistic regressions to jointly explore the influence of selected socio-demographic and health care factors on caregiver responses to fever in their young child and the clustering of these behaviours at the village and health unit catchment area levels. Only a few studies have examined the determinants of these behaviours using multivariate regression and this is the first study to document the existence of clustering in caregiver fever management practices. The data collected as part of the C-IMCI survey presented an excellent opportunity to undertake this comprehensive analysis due to the large sample size, the recruitment of caregivers and children in their homes rather than at health facilities, the use of actual, recent fever episodes, and the cluster nature of the survey design. This discussion will comment on the interpretations of the observed associations and clustering followed by an examination of potential implications for home and community programs and study limitations.

3.5.1 Interpretation of factors associated with caregiver fever management practices

Socio-demographic factors: Of the socio-demographic factors studied here, statistically significant associations were found between fever management practices and child age, caregiver education level, household size and region of residence and these associations varied across behaviours. The lack of influence of child sex and caregiver age on the management of fever has been reported in studies in Tanzania and Kenya^{83;95}. The finding that children aged three to eleven months are more likely to be treated at home with an antimalarial corresponds relatively well with parasitological findings and indicates that those children most likely to receive an antimalarial in the home are those at

greatest risk for malaria. In the Gambia, children aged six to eleven months had the highest parasite densities and highest rates of probable malaria illness compared to all other under-five age categories²³. An association between child's age and care seeking was not found, however the small age range of children in this study means that such an association cannot be excluded^{83;94}.

The reasons underlying the positive association between household size and home treatment with antimalarial are unclear, but there are several possible explanations. Caregivers living in larger households are generally responsible for more children and as a result, may have additional experience with childhood fever, greater confidence in their ability to treat it in the home and better awareness of the rapid antipyretic benefits of administering an antimalarial such as chloroquine¹¹⁵. Initial home treatment may also be facilitated by the presence of drugs leftover from other illnesses in the family, which would be expected to be more prevalent in larger households. Larger household size could also reflect other underlying differences, such as higher household income or the presence of other support structures, which cannot be explored in this analysis.

Caregiver education level was an important correlate of home treatment with an antimalarial, with primary education or beyond associated with a nearly two-fold increase in the likelihood that a child received an antimalarial after adjustment for other factors. Conversely, education level did not exert a strong effect on care seeking behaviours, although there was some evidence that further education was associated with an increased likelihood of seeking outside care in the multivariate model (OR = 1.29; 95% CI: 0.85, 1.98). This weaker association of education to care seeking behaviour is not due to the inclusion of previous home treatment practices in the care seeking models, since the

association is even weaker when this variable is excluded from the model (results not shown). This appears to contrast a study in Tanzania that found higher maternal education was associated with increased promptness of care seeking, however, this study did not adjust for other factors and did not find a large difference (70% of those with seven or more years of education would take their child within hours compared to 61% of those with less than seven years)⁹³. These differences in the influence of education across caregiver responses to fever require further study and validation.

The mechanisms by which caregiver level of formal education exerts its influence on fever management practices are not clear, but there is some evidence to suggest it could be at least partly mediated through improved knowledge. A study in Tanzania reported that caregivers of children with primary or better education levels had significantly better knowledge of use of antimalarials for fever¹¹⁵. Higher education could also be acting as a proxy for other variables not measured as part of the C-IMCI survey. For example, both income level and urban residence are correlated with education levels and have been shown in other studies to be associated with use of antimalarials in the home and care seeking behaviour^{95;116}. With these factors uncontrolled for in the present study, it is difficult to speculate on how the observed associations might be modified.

Region of residence emerged as a strong independent predictor of care seeking behaviour, however these findings differ qualitatively from those reported in a bivariate analysis of the Uganda DHS conducted over a similar time period⁸⁹. The UDHS study found caregivers in the central region most likely to seek care while caregivers in the western region were least likely⁸⁹. In contrast, this study found caregivers in the central

and western regions to be similar and caregivers in eastern districts and northern regions to be significantly more likely to seek care and to do so within 48 hours of fever onset. These differences are probably due to the small number of districts available in this study that may not have been representative of their region as a whole. However, the existence of regional differences in care seeking behaviours after adjustment for several covariates suggests that region or district specific factors likely play an important role. Such factors could include disparities in the quality of health care services, economic conditions, and varying social and cultural practices and health beliefs. Further study will be required to explore reasons behind these striking differences in care seeking behaviour. The lack of an association between region of residence and antimalarial administration in the multivariate model indicates that regional variation can be largely explained by differences in other underlying determinants.

The C-IMCI study did not collect information on other potentially important socio-demographic correlates of caregiver responses to fever. An important omission is a measure of income or wealth⁹⁵. In Uganda, user fees at public health facilities were removed in 2001, but this is unlikely to have adequately addressed the economic burden associated with care seeking because user fees account for only a small proportion of care seeking costs^{117;118}. Although traditionally difficult to measure in poor populations, simple methods are now available that document the presence or absence of a small number of household items (i.e. bicycles, roofing materials, radios) and create an index. These straightforward additions to survey questionnaires can provide valuable information for both understanding factors and exploring differences even within poor populations in terms of intervention coverage^{95;119}.

Health care factors: Of the health care aspects studied, geographic access emerged as an influential factor, but only for care seeking behaviours. As noted in numerous other studies, physical distance from a source of care is a key determinant both of the decision by caregivers to seek care for their child and to do so promptly^{91;93;94;120}. In many parts of Uganda where individuals must walk to access care, and monetary and time costs rise with increasing distance, even small distances can act as a barrier and contribute to delays in care seeking. A two-fold increase in likelihood of seeking outside care was observed among caregivers living within five km of the nearest care source. A recent study in the Kabale district in south-western Uganda found a significantly increased likelihood of attending a health facility within one day of symptom onset (OR=2.58; 95% CI: 1.29-6.41) among those living closer to a health facility than a shop⁹².

The independence of home use of antimalarials for childhood fevers from distance to a health care source is not surprising when viewed in light of the many studies showing that most caregivers obtain antimalarials from local shops, leftover supplies from previous episodes, and other family members or neighbours rather than health facilities^{83;87;91;121}. This finding is supported by a study in Makueni district of Kenya that found no association between distance from nearest health centre and presence of chloroquine in the blood of 317 under-five children⁷⁵. The remarkably easier access to antimalarials and other western drugs through informal sources was demonstrated in Kenya by mapping of households in relation to pharmacies, shops, and health facilities. Nearly 90% of rural homesteads were within one km of a shop or pharmacy while only 32% were within two km of a formal health facility⁸³. Urban residents lived within one km of all facility types, but still preferred shops since they also cater to other consumer

preferences including longer and more convenient hours of operation, friendly and fast service, and provision of amounts according to what a customers can afford⁸³.

This study did not find any significant associations between caregiver fever management practices and perceived quality of care and extent of community outreach by village health workers. There was some evidence of a trend towards increased likelihood of care seeking among those with a positive perception of health services in the univariate model (OR=1.28; 95% CI: 0.93, 1.76), but this effect was reduced after adjustment for covariates. The initial trend towards increased care seeking among those reporting a visit of a health care worker to their village or household within the last five months (OR=1.41; $p=0.12$) was eliminated after adjustment for covariates, suggesting it was a spurious finding as a result of its relationship with distance to a health facility, such that those living nearer a health facility would be most likely to receive a visit. The limited information regarding health services, and particularly objective evaluations of quality such as availability of drugs, diagnostic services and adequacy of physical examination procedures, mean that a greater influence of health care aspects on caregiver fever management practices cannot be ruled out. These results do suggest, however, that any influence of health care variables would be stronger on care seeking behaviours than on initial home treatment behaviours.

Linkage between home treatment and care seeking behaviours: A key finding of this study was that previous home treatment was the strongest correlate of seeking outside care for fever and for seeking care promptly, demonstrating a strong linkage between initial home treatment and care seeking behaviours after adjusting for other influences. Caregivers treating at home with an antimalarial were significantly less likely to seek care

and to seek care within 48 hours compared to caregivers who gave no initial home treatment. Delays in care seeking among those treating first at home were also noted in a facility-based study in the Tororo and Busia districts of Uganda, where the median duration of fever among children until time of presentation was nearly twice as long for those given some form of home treatment compared to those for whom nothing was done (3.99 versus 2.16 days, respectively)⁸⁷. A study of caregivers in Bungoma District Kenya reported no significant difference in the mean time to care between those providing an antimalarial in the home and those first attending a health facility (mean delay of 3.2 and 2.7 days, respectively), however the direction of effect was similar and results may have corresponded to those reported here if presented categorically as in the present study⁸⁵. Of concern for achieving prompt treatment of fever with an antimalarial was the finding that home treatment practices that did not include an antimalarial were also associated with a decreased likelihood of seeking care and increased delays of a magnitude similar to that for antimalarial administration. These children are at risk of not receiving an antimalarial for fever or of receiving one too late to avoid consequences.

The cross-sectional nature of the data and the lack of information on potentially important covariates such as severity of fever, presence of other symptoms, etc., preclude conclusions regarding the direction of association between home treatment and care seeking behaviours. However, qualitative studies support that home treatment may delay presentation to a health facility as caregivers will wait to see whether any initial treatment has been effective before expending additional resources to seek care, particularly if the fever is not severe⁸². For resource poor populations, home treatment is an attractive initial response. Economic studies document considerable savings in time and monetary

expenditure^{118;122}. In Burkina Faso, the out-of-pocket expenditures for seeking care at a health facility were twice that for home treatment¹¹⁸. Time savings associated with home treatment were highlighted in a study in Ghana which found that time costs to purchase drugs at shops averaged between one and two minutes, while average travel time to a health facility ranged between 16 and 24 minutes and total time spent at facilities from 112 to 250 minutes¹¹⁷.

3.5.2 Interpretation of clustering of behaviours

The second focus of this study was to explore the clustering of behaviours at the village and health unit level. The presence and magnitude of clustering was measured using the pairwise odds ratio, which expresses the dependency of behaviours among caregivers belonging to a given village or health unit. There is evidence of moderate clustering at the village level, but little evidence of residual clustering at the health unit level once clustering at the village level has been accounted for. This trend for decreased clustering with higher levels of aggregation was also observed in the studies of diarrhoea occurrence in low income countries and narcotic drug use in the United States and is a common pattern observed in estimates of intracluster correlation^{111;112;123;124}. The magnitude of village level clustering observed here is similar to the range of those reported for clustering of fever prevalence at the village level in Zambia (1.04), Indonesia (1.15), Malawi (1.18) and Nepal (1.34) and for marijuana use in neighbourhoods in the United States (1.3 for reports of ever using marijuana and 2.0 for recent person-to-person sharing of marijuana)^{97;123}. The selection of consecutive households at the village level and of parishes closest to the health centre may have overestimated the degree of behavioural dependency since these caregivers may have more in common and more

contact with each other than caregivers selected randomly throughout the village and health centre catchment areas. In addition, village PWORs were based on data from only 72% of villages for home treatment and care seeking and 60% for promptness of care seeking, resulting in imprecise estimates and uncertainty regarding the representativeness of these values should villages unable to contribute to the PWOR estimates differ from those included.

The greater degree of village level clustering observed for care seeking behaviours in comparison to home treatment could reflect a larger role of family and community members in decisions related to seeking care outside the home. A study of intra-household decision making for childhood fever management in Kenya reported significant and complex community interaction in the decision to seek care and that mothers reported seeking more advice and financial assistance for health care visits than home care¹²⁵. This is not unexpected, since as noted, attending a health facility usually involves additional time and out-of-pocket costs than home treatment and therefore additional input, financial assistance, and permission is often required or sought out from other family members and community members^{117;118;125}. Additional research is needed to further explore the reasons behind these observed differences in behavioural dependency across fever management practices.

The persistence of clustering at the village level after adjustment for several socio-demographic and health care variables argues for further contemplation of the potential factors contributing to this clustering, particularly for promptness of care seeking. The greater clustering at the health centre level for home treatment behaviours but not care seeking is also intriguing. In this regard, attempts should be made to explore the

influence of community level factors which may influence individual level behaviours and thus contribute to the observed clustering. Factors may include village level experience with malaria, level of wealth at the cluster level and other factors yet uncovered⁹⁰.

3.5.3 Implications for home and community programs

These results have important implications for home and community programs that aim to decrease child mortality due to malaria and other febrile illnesses. Uganda and several other sub-Saharan countries are rapidly scaling up programs for the home management of fever. These programs encourage caregivers to treat uncomplicated fevers with antimalarials administered in pre-packaged doses, which are distributed as near to the home as possible, through drug shops, community volunteers and other venues. Initial evaluations of such programs in Ethiopia and Burkina Faso have found an increase in the proportion of children with fever receiving early treatment with an effective antimalarial and substantial reductions in child mortality and morbidity^{44;45}. The strong link between initial home treatment and care seeking behaviour noted in this study underscores the need to consider these behaviours jointly as intervening on one will affect the other and also argues for rigorous evaluation alongside scale-up efforts. Home management programs focussed on malaria alone could lead to delays in care seeking and negatively affect outcomes of other illnesses that cause fever. In malarious areas, approximately 20 to 60% of childhood fevers are attributable to malaria, leaving a good proportion that may require additional treatment for other conditions³². Of particular importance are acute respiratory infections which do not respond to chloroquine but present similar symptoms (i.e. fever, cough, difficulty breathing) and are difficult to

distinguish from malaria in the absence of laboratory results^{11;12}. Acute respiratory infections account for a considerable proportion of child morbidity and mortality in sub-Saharan Africa and should not be neglected in the eagerness to address malaria.

Within the home and community based approach, strategies and educational messages should focus on comprehensive fever management and encompass the major causes of serious febrile illness. Careful attention should be paid to the messages given to caretakers, focussing on the importance of continual monitoring for danger signs and lack of improvement and of seeking care immediately in these cases. Interventions to influence care seeking behaviours will also need to address physical access and region specific factors. The interdependency of caregiver behaviour emphasizes the importance of extending programs beyond the primary caregiver and addressing the broader community. There are often strong community norms about what constitutes correct management of fever and other childhood symptoms and this influence could be harnessed to promote positive behaviours and increase compliance to recommended fever management practices. Further study is recommended as a priority to determine the influence of home treatment programs for malaria on other causes of fever, especially acute respiratory infections.

3.5.4 Limitations

The cross-sectional nature of the present survey prevents examination of the direction of effect for the associations reported here. All data were based on caregiver self-report using a non-validated survey tool and are vulnerable to report and recall bias. The C-IMCI interviews were conducted by non-medical personnel, which minimizes the tendency for caregivers to feel pressure to report taking their sick child to care or to

misrepresent their use of antimalarials⁷⁵. However, should certain groups such as those with higher education or from certain districts have been differentially inclined to respond in a socially desirable way, an effect on the associations reported here cannot be excluded. Caregivers were asked to recall fever symptoms and subsequent case management actions over a relatively short recall period of two weeks, a period which has been shown to elicit acceptable, although not optimal results^{126;127}.

A certain amount of misclassification of fever, fever management practices and socio-demographic and health care covariates is inevitable, however, this misclassification is expected to be non-differential and underestimate the associations observed here¹²⁸. The ascertainment of childhood fever cases relied on caregiver reports of fever and sub-study I demonstrated that caregivers will rarely miss a fever, however some over-reporting of fever cannot be excluded. Fever cases were not restricted to completed episodes and as a result, unresolved fever episodes may have been misclassified as not seeking care or not providing home treatment even though one or more of these activities might occur over the illness course. However, these are expected to represent only a small proportion of fever cases. Several independent variables, including caregiver and child age, caregiver education level and distance to care source were also susceptible to misclassification; however the broad groupings would have minimized any impact.

The potential for selection bias due to refusal or non response is considered low. Information was not available to determine the overall survey response rate. However, a significant number of caregiver refusals is unlikely based on experience with household surveys in similar populations and the enthusiasm for participation in areas where the

author observed data collection procedures^{95;129}. Response rates for individual questions were quite high and approximately 80% of cases were available for each outcome even in the regression analysis that included only those with complete data for all variables. In addition, excluded cases were similar to included cases with regard to the covariates of interest.

The C-IMCI survey did not collect information on a number of potentially important determinants of maternal responses to childhood fever. In addition to the socio-demographic and health care aspects noted previously, data were lacking on severity of fever, concurrent presentation of additional symptoms, health beliefs and level of knowledge regarding malaria, and intrahousehold factors such as the need for permission from child's father or other family member with decision-making responsibility^{94;95;130;131}. The importance of these factors among caregivers interviewed as part of this survey as well as how they may modify or confound the associations reported in this study is unclear and an area worthy of further exploration.

The C-IMCI survey employed a complex design incorporating sampling procedures that deviate from simple random sampling upon which the principles of statistical analysis are based. The method of alternating logistic regressions was used to address within-cluster correlation over other potential approaches, which include jackknife and bootstrap techniques, variance correction algorithms and random effects models (multi-level analysis)¹¹⁴. Bootstrap and jackknife applications require detailed information on sampling strata and primary sampling units for the calculation of weights and adjustments of variance and this information was not available for the current study. The multi-level analysis algorithm performs regressions for each cluster and requires sample sizes of at

least 25 respondents in each cluster to be valid¹³². The small numbers available at the village (6 or less) and the health unit (17 or less) for the regressions precluded the use of random effects models.

Adjustment for clustering using alternating logistic regressions analysis addresses only part of the design issues associated with the C-IMCI survey. The lack of proportional-to-size sampling at the village level and sample weights result in a sample that may not be representative of the target population. A comparison of the C-IMCI sample and the nationally representative UDHS survey conducted over a similar time period found very similar socio-demographic and fever management profiles suggesting a relatively representative sample was achieved despite the methodological shortcomings⁸⁹. The need for sample representativeness is not as crucial for analyses which explore relationships among variables than for those that aim to provide point estimates of behaviours.

The generalization of these results beyond the study population should be undertaken cautiously. The associations obtained here were based on caregivers of children less than two years old sampled from ten districts in Uganda among parishes located nearest to the government health facility. Consequently, their applicability to caregivers and communities located further away or to other areas within or beyond Uganda is uncertain. The relative importance of determinants and the degree of clustering for caregiver fever management practices is likely to vary across social, cultural, economic and environmental settings and the collection of local data is important to tailor programs appropriately.

4.0 SUB-STUDY III: PROPOSAL DEVELOPMENT GUIDELINES AND SAMPLE SIZE FRAMEWORK FOR THE HCMMP STUDY

4.1 BACKGROUND

The Home and Community Management of Malaria and Pneumonia (HCMMP) study is an initiative to develop, implement and evaluate the effectiveness of a home and community based program for fever in under-five children living where both diseases are endemic. This collaborative project is being coordinated by the HCMMP working group, an interagency working group comprised of researchers from the Department of Epidemiology and Community Medicine at the University of Ottawa, the Division of International Health of the Karolinska Institute, UNICEF and Tropical Disease Research (TDR) of the WHO. The HCMMP program will be implemented and evaluated in several countries following a common study protocol. This sub-study describes the development of a set of proposal development guidelines and presents an analysis of sample size and power requirements for the HCMMP study.

4.1.1 Overview of the HCMMP study

Malaria and pneumonia are leading causes of morbidity and mortality among children less than five years old in sub-Saharan Africa¹³³. These two diseases resemble each other clinically such that they are indistinguishable even by trained health staff and can kill young children rapidly, often within two days of symptom onset^{10;26;134}. Most deaths result from lack of early diagnosis and prompt treatment with an effective antimalarial and/or antibiotic^{4;135}. The HCMMP study aims to compare the safety and effectiveness of three methods of providing care for common childhood illnesses - one that provides home and community based care for malaria and pneumonia together, one that provides home and community based care for malaria alone, and one that is

representative of normal care and based primarily at the health facility. Several countries in sub-Saharan Africa are in the process of scaling up home and community initiatives to presumptively treat childhood fevers with antimalarials based on limited evidence. Replication in a variety of settings is crucial to support such extensive policy and programmatic changes^{99;136}. As well, the clinical and epidemiologic overlap between malaria and pneumonia is well-established and there is growing concern that the scale-up of programs that target only malaria treatment for fever could lead to important delays in obtaining treatment for pneumonia¹⁰⁻¹². Thus, evidence on the impact of addressing both conditions at the community level and any potential benefits over addressing that of malaria alone is needed to provide guidance for countries to adequately define the scope of their programmes and home and community management policies.

The HCMMP program consists of an integrated set of home and community based strategies that address the complex array of factors contributing to child mortality and integrate the involvement of the wide range of stakeholders (caregivers, drug vendors, traditional healers and formal medical staff) who play a role in the management of these common childhood illnesses. The aim is to modify behaviours towards: early suspicion of illness onset, particularly fever; prompt administration of fully efficacious drugs; complete adherence to treatment dosing and schedule; early identification of severe illness, complications or lack of improvement; and prompt referral to a functioning health facility. Four programmatic elements will be established to promote these modifications: 1) Information, Education, Communication (IEC) to families and communities; 2) an effective and sustainable study drug supply system; 3) training of and skills support for study drug distributors and caregivers; and 4) an integrated referral and monitoring

system. These components are based on experiences of earlier investigations into home and community based strategies for malaria and pneumonia. By targeting multiple levels including the individual, the community, and the health system, these components have the potential to achieve a synergistic impact on child survival^{44;45}. In the malaria only arm, communities will receive these four elements as they relate to malaria alone (i.e. community access to antimalarials but antibiotic access through health facilities alone). In the combined arm, these four elements will address both malaria and pneumonia and community access to antimalarials and antibiotics will be established.

The HCMMP program will be evaluated in a multicentre, cluster-randomized controlled trial. Communities will be randomly allocated to one of three arms: home and community management of malaria and pneumonia, home and community management of malaria only, and comparison group. This design provides direct, country level information regarding both the effectiveness of malaria home and community management and the combined approach in comparison to the care that is provided in line with national practices. In addition, it allows an indirect assessment of the potential additional benefits of addressing malaria and pneumonia together compared to malaria alone through the pooling of results across study sites. Implementation will focus on areas with poor access (financial and/or geographical) to adequate health care services and medication, high rates of child mortality attributable to malaria and pneumonia and where presumptive treatment for both conditions is the treatment policy. The primary outcome will be a reduction in overall under-five mortality. Secondary outcomes include cause-specific mortality and severe morbidity, adherence to recommended antimalarial

and antibiotic treatment regimens, intervention coverage, drug-related adverse events, and antimicrobial resistance levels.

4.1.2 Development of proposal development guidelines and sample size framework for the HCMMP study

At the initial design phase during which this sub-study was conducted, two primary needs for the planning and development of the HCMMP study were identified. The first need was for a set of proposal development guidelines to help interested country teams plan studies addressing key methodological issues in a consistent manner and to facilitate the selection of research partners by the TDR Steering Committee on Implementation Research. One of the key questions the HCMMP study seeks to address is whether a combined approach addressing both malaria and pneumonia is more effective than addressing malaria alone. No single site will be able to support a large enough study on its own so results will be pooled across a number of sites. This combination of results requires a certain degree of comparability and consistency in terms of program elements, study population, and a common research methodology.

The second identified need was for a comprehensive framework to assess anticipated sample size and power requirements. The HCMMP program involves a wide range of community members, including families, shopkeepers, village leaders, traditional healers and formal health care workers, and precludes randomization at the individual level. Randomization at a group, or cluster level, has profound implications on all aspects of study design, and particularly sample size requirements¹²⁴. To account for the randomization of groups rather than individuals, the sample size must be increased¹²⁴. The degree of increase is determined largely by the size of clusters and the extent to which cluster members behave more similarly than those of the general population¹²⁴.

Methodological reviews of cluster trials indicate that few investigators adequately account for the presence of clustering in sample size calculations and often undertake trials that are poorly designed and underpowered^{137;138}. This is in large part due to the greater complexity of calculations and the poor availability of critical information such as the degree of intracluster correlation and cluster sizes¹²⁴. Determining a sample size that delivers sufficient power and is at the same time feasible is particularly challenging for cluster trials of complex interventions. Small changes in sample size parameters tend to have a large impact on the required study size¹²⁴. Sample size and power sensitivity analyses which vary sample size parameters across a plausible range of expected values can help anticipate these and are highly recommended early in the design stage¹²⁴.

4.2 OBJECTIVES

The objectives of this sub-study are i) to develop in collaboration with potential research partners a set of proposal development guidelines to ensure comparable methods across study sites and ii) to provide a framework for evaluating sample size and power requirements for the HCMMP study.

4.3 METHODS

4.3.1 Development of the proposal guidelines

The guidelines were developed in two stages. The first stage involved consultations with program and policy stakeholders. Initial meetings of the HCMMP working group identified several elements that would require early, meaningful input from country level representatives to ensure results would be relevant to child health policies at international, national and regional levels. These factors included principal research questions, study design, comparison group composition, and sample size parameters, among others. To

obtain this input, a meeting of policy experts and researchers from malaria endemic countries was organized jointly by the WHO, UNICEF, the University of Ottawa and the Karolinska Institute and held in Entebbe, Uganda 11-15 February, 2002.

The purpose of this meeting was to bring together researchers and policy makers to share their experience and expertise in community level interventions for malaria and pneumonia and to reach consensus on the relevance of an integrated approach and related study design issues. Approximately 30 individuals were invited including members of the HCMMP working group, representatives from Burkina Faso, Ethiopia, Ghana, Nigeria, and Uganda, and selected international experts. Participants were identified in consultation with the WHO and included individuals from Ministries of Health, research institutes, and universities. African representatives contributed considerable practical experience with home management research and implementation, having been involved in recently completed TDR-funded pilot studies (in Burkina Faso, Ghana, Nigeria and Uganda) to explore home management of fever. The format of the meeting was designed to facilitate information sharing and consensus building around the identified protocol issues and consisted of presentations of relevant background material, interactive round table discussions, and group work. Representatives discussed protocol issues in detail in working groups and prepared recommendations that were taken to the plenary for consensus.

The second stage built upon this preliminary framework to create and finalize the proposal development guidelines that would be sent to interested country teams. Within the guidelines, attempts were made to balance two competing goals: 1) achieving a certain level of consistency and comparability across studies, and 2) allowing for

flexibility and innovation. To achieve the first goal, elements of the study design and the intervention itself that would need to be sufficiently similar to facilitate the pooling of results through meta-analysis were fixed in the guidelines. To achieve the second goal and to recognize the uniqueness of each country/regional situation and research team experience and expertise, allowances were made for a considerable amount of creativity in development of the intervention materials and drug delivery systems, in selecting secondary outcome measures and numerous other methodological decisions. The guidelines were formatted to follow the guidelines for randomized controlled trials developed by the Canadian Institute for Health Research (CIHR) as much as possible. Examples of these guidelines are available online at: http://www.cihr-irsc.gc.ca/services/funding/apply/instr/clinicaltrials_e.shtml. They are also in line with the CONSORT guidelines for randomized clinical trials that have recently been adapted for cluster trials¹³⁹.

An initial draft of the proposal guidelines was prepared and taken for review and finalization to a meeting held at the Karolinska Institute in Stockholm from June 13-14, 2002. Participants in this meeting included members of the HCMMP working group (Dr. George Wells, Dr. Jane Kengeya-Kayondo, Dr. Melba Gomes, Dr. Stefan Peterson, Ms. Karin Källander and the author) and the coordinator for the recently formed home management of malaria and pneumonia network (Dr. Betty Mpeka, Ministry of Health Uganda). The feedback and comments on the draft guidelines were incorporated in the final version of the guidelines.

4.3.2 Framework for sample size and power

The primary outcome of the HCMMP study is the reduction in overall under-five mortality rates. Mortality will be measured by following up a cohort of study children over the two year study period and the rate in each group will be calculated by dividing the number of child-deaths over the number of child-years of follow-up.

Sample size formula: Sample size estimates were calculated using the formula for event rates under conditions of cluster sampling given by Donner and Klar¹²⁴:

$$k = 1 + \frac{(Z_{\alpha/2} + Z_{\beta})^2 (\lambda_1 + \lambda_2)}{t(\lambda_1 - \lambda_2)^2} \times IF_t$$

where:

$$IF_t = 1 + \frac{CV^2(\lambda_1^2 + \lambda_2^2)t}{\lambda_1 + \lambda_2}$$

The number of clusters required per treatment group is represented by k and α is the significance level and β is set risk for type two error. The average under-five mortality rate for the intervention group and comparison groups are given by λ_1 and λ_2 , respectively. The child-years of observation in each cluster are t . The impact of clustering is taken into account by calculating the inflation factor (IF), or design effect, which represents the increase in sample size over that of individual randomization for a given set of parameters. The CV value represents the coefficient of variation for the between-cluster variation in mortality rates and larger values are associated with higher inflation of the sample size¹⁴⁰. The addition of an extra cluster ($1 + k$) to each study arm is recommended when the number of clusters is small to offset the use of critical values based on the normal distribution when the t-distribution is used in the analysis^{124;140}. The formula used to calculate k is for a completely randomized design involving three parallel

study groups. This approach is conservative should matching or stratification, which act to decrease the coefficient of variation, be employed¹²⁴.

The overall alpha level was set to 0.05 and the desired study power to 0.80. A Bonferroni correction was applied to the alpha level ($0.05/2=0.025$) to account for the additional comparison made using the three-arm design. The number of clusters required for the entire study was calculated as $3k$. The value for k was based on the number of clusters required in the malaria versus the comparison arm since the impact of the intervention addressing malaria alone is expected to be less than that for the intervention addressing both conditions. The value of k was rounded upwards to the nearest whole number.

Estimation of sample size parameters: A sensitivity and power ($1-\beta$) analysis was performed by varying sample size parameters based on ranges obtained through a combination of consultations with country level researchers and literature searches. One of the key parameters to establish is the degree to which participants from the same cluster are more alike than would be the case in a randomly selected sample, accounted for by CV in the formula. The effect of CV on sample size estimates is equivalent to that of the intraclass correlation coefficient (ICC), which expresses the percentage of total variation that is accounted for by between-cluster variation¹⁴⁰. The CV for between-cluster variation is calculated by dividing the standard deviation by the average mortality rate¹⁴⁰. For most health outcomes, CV values are less than 0.25 and few exceed 0.50, the latter which corresponds to a range in mortality rates of 0 to 80 per 1,000 child-years¹⁴⁰. Obtaining an appropriate estimate of CV prior to the start of the trial is notoriously difficult. Few studies publish CV values and those that are reported may be imprecise

due to the small number of clusters involved¹²⁴. In addition, CV values are inversely related to cluster size and depend on study design (i.e. smaller for matched than completely randomized)¹²⁴. There is little information regarding their generalizability, therefore local data are highly preferred¹²⁴. For the current analysis, relevant estimates for CV for under-five mortality rates were sought in the published literature from studies conducted in sub-Saharan Africa either from cluster trials where communities were randomized or where mortality surveillance was conducted and rates reported at the community level. When data were adequate, estimates of CV were calculated using the equation suggested by Hayes and Bennett¹⁴⁰, which adjusts the value calculated from sample data for the effects of sampling error:

$$CV_{\text{est}} = \sigma c^2 / \lambda \quad \text{where } \sigma c^2 = s^2 - \lambda \text{ Av}(1/t_j)$$

In this equation, λ is the overall mortality rate across clusters, s^2 is the variance of cluster mortality rates, and $\text{Av}(1/t_j)$ is the average of the reciprocal of the number of child-years (t_j) of follow-up in the j th cluster¹⁴⁰.

Information on other key parameters, including anticipated cluster sizes, baseline under-five mortality rates (BMRs), and estimates of the minimally clinically important difference, or effect size (ES), were obtained from African researchers at the consultative meeting in Entebbe. A Delphi exercise was performed to determine the ES for each comparison¹⁴¹. All Entebbe meeting participants working in malaria endemic areas were asked to submit their anonymous opinions of what constitutes a significant reduction in under-five mortality from malaria alone, malaria and pneumonia and the difference between malaria and malaria and pneumonia combined. A mean value was determined for each of these and used in subsequent calculations. Participants were also asked to

provide detailed information on possible randomization units and their population size and proportion of children under-five.

Adjustment for losses to follow-up: The impact of potential losses to follow-up on sample size and power at both an individual level and a cluster level were explored. Losses to follow-up in a cohort of children living in rural communities followed for two years are expected to be at least 20%¹⁴². These losses will reduce the overall cluster size by decreasing the number of child-years of observation. Such losses were adjusted for by reducing the overall number of child-years by 20% and by 30% over the two year study period and calculating the number of clusters required given the revised size in child-years assuming an annual census. It was assumed that losses would occur similarly each year (i.e. 10% or 15% per year) and that children would be lost to follow-up at the mid-point of each year and so still contribute 0.5 years. For example, given an initial cohort of 100 children, 180 child-years of follow-up would be available over the two year study period compared to 200 child-years should all children remain in the study (Year 1: 90 + 10 x 0.5 years contributes 95 child-years plus Year 2: 80 + 10 x 0.5 contributes 85 child-years). The impact should one or more entire clusters drop out of the study was considered in a power analysis.

All calculations were performed in Microsoft EXCEL[®]. Results are presented mainly in graphic form to illustrate the trends in sample size and power estimates across varying levels of parameter inputs.

4.4 RESULTS

4.4.1 Proposal development guidelines

The meeting in Entebbe was well-attended and consensus was reached on many research question and design parameters for the HCMMP study. Initial proposal guidelines based on these discussions were modified slightly at the Stockholm meeting, but continued to follow the consensus of the Entebbe meeting. A summary of the meeting results are included in Appendix IIIA and the final proposal development guidelines are attached as Appendix IIIB.

4.4.2 Sample size and power sensitivity analysis

4.4.2.1 *Estimation of sample size parameters*

Coefficient of variation: The literature search uncovered one published estimate of CV for overall under-five mortality and three studies reporting cluster-specific mortality rates. A CV value of 0.29 was obtained from the unmatched insecticide treated net (ITN) intervention in Mwanza, Tanzania with a primary outcome of overall mortality in children aged 1 to 59 months collected across 51 'zones' comprised of approximately 200 children^{140;143}. One ITN trial in The Gambia, which grouped villages into five administrative areas of approximately 1,000 to 3,000 children based on cultural and ecological variation, provided sufficient information to calculate a CV¹⁴⁴. Based on data collected during the post-intervention year, there were 183 deaths over 10,407 child years in the control group for an overall mortality rate in the five areas of 17.6/1,000 (range of 9.9 to 19.2/1,000) and a CV of 0.12. Greater variability was noted in the intervention due to uneven effect across clusters; overall mortality rate of 14.4/1,000 child-years (range of 3.8 to 21.4/1,000) for a CV of 0.46.

The wide range in cluster specific rates of under-five mortality in the remaining two studies in Ethiopia and Burkina Faso correspond to large unmatched CV values similar to, or higher than, CV values as observed in the Tanzania and The Gambia ITN trials^{45;145}. For example, among villages of 100 or more children (range 104-405) included in a demographic surveillance program in Burkina Faso, overall under-five mortality rates per 1,000 child-years ranged from 5.0 to 86.2 (1993); 15.2 to 120.0 (1994); 6.6 to 77.8 (1995); 5.8 to 138.5 (1996); 0.0 to 131.8 (1997); and 6.6 to 80.5 (1998)¹⁴⁵. These surveillance data also demonstrates that under-five mortality rates observed in one year do not reliably predict those in subsequent years, with correlations ranging from a low of 0.27 (1995 and 1996) to a high of 0.65 (1997 and 1998). Furthermore, the relative ranking of each village was not consistent from year to year, with correlations lower than those observed for mortality rates¹⁴⁵.

Effect size: At the Entebbe meeting, twenty-one attendees participated in the Delphi exercise to determine the ES of importance for the malaria versus comparison group, the malaria and pneumonia versus comparison group and the malaria versus the combined group. The percentage reduction of importance for malaria alone compared to the comparison group ranged from 6.3% to 40.0%, with a mean of 24.8%. Values for the combined approach versus the comparison group ranged from 20.0% to 60.0% with a mean of 39.7%, and for the combined approach versus malaria alone from 7.0% to 37.5% with a mean of 19.8%.

Cluster size: Information on administrative structures and an estimation of the population comprised of children under-five years of age per cluster obtained from potential study sites at the Entebbe meeting is summarized in Table 4.1. Administrative

constraints in some countries may limit options for unit of randomization to the sub-county or equivalent level. For example, in Uganda, health centres operate at a sub-county level making it difficult logistically and ethically to assign villages or parishes belonging to the same health centre to different study arms without increasing the possibility of contamination or complicating the delivery of the intervention, which seeks to actively involve health centre staff. Similarly, in Nigeria and Burkina Faso community level units correspond with community clinic catchment areas.

Table 4.1 Estimated under-five population estimates for selected countries by administrative level

Administrative Level	Burkina Faso	Ghana	Nigeria	Uganda
District/LGA [§]	30,000	30,000	60,000	54,000
County/Ward	--	7,000	6,000	12,600
Sub-county/Community	2,000	--	2,000	3,600
Parish	--	--	--	900
Village	400	400	330	90

[§]- Local Government Area

4.4.2.2 Sensitivity analysis

Based on the information outlined above, baseline under-five mortality rates were varied from 20/1,000 to 100/1,000 child-years in units of 20/1,000 child-years; cluster size was varied from 100 to 1,000 in units of 100; ES for malaria versus comparison was varied from 0.20 to 0.30 in units of 5; and CV values were varied from 0.15 to 0.50.

Results of the sample size and power sensitivity analyses are summarized in a series of graphs (Figures 4.1 to 4.7). A mid-range set of parameters was selected to show the pattern of influence of input parameters on sample size requirements, whereby CV = 0.25, ES = 0.25, BMR = 80/1,000 and varied accordingly. Table 4.2 presents sample size

estimates in terms of number of clusters, number of children and corresponding inflation factor for this range.

Table 4.2 Sample size requirements for a two year, three-arm cluster randomized trial to reduce under-five mortality in the malaria arm by 25% compared to the control arm, assuming a baseline mortality rate of 80/1,000 live births and a CV of 0.25

Cluster size (#U5)	No. of clusters per arm	Total number of clusters	Total number of children	Inflation factor [§]
100	33	99	9,900	2.0
200	25	75	15,000	3.0
300	22	66	19,800	4.0
400	20	60	24,000	4.8
500	20	60	30,000	6.0
600	19	57	34,200	6.9
700	19	57	39,900	8.0
800	18	54	43,200	8.7
900	18	54	48,600	9.8
1,000	18	54	54,000	10.8

[§] - compared to the 4,980 children required under individual randomization

The trade-off in number of clusters required by cluster size is evident in Figure 4.1; the required number of clusters falls with increasing cluster size until cluster size reaches approximately 400 to 500 children. Additional increases in cluster size offer minimal reductions in cluster requirements but significantly increase the overall sample size. This pattern is reflected across input parameters as denoted in Figure 4.2.

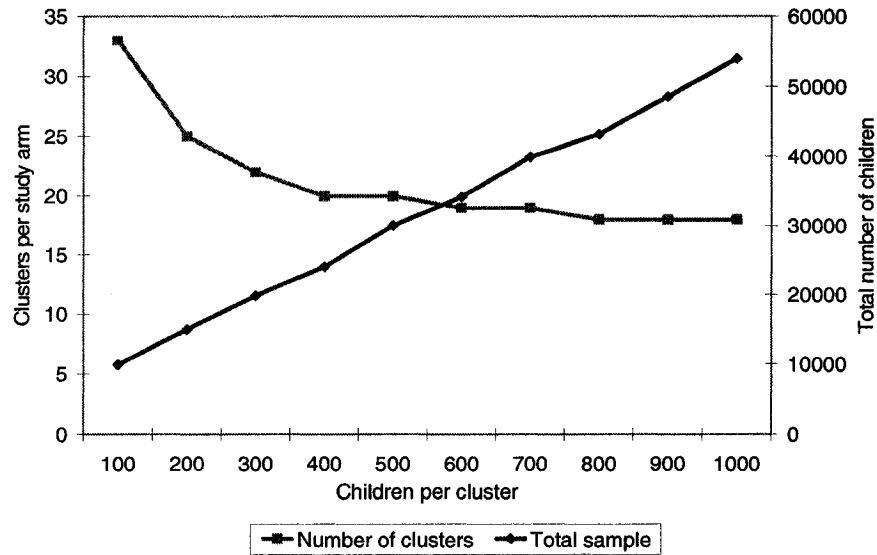


Figure 4.1 Total number of children and number of clusters per study arm by cluster size assuming baseline mortality rate of 80/1,000 child-years, a coefficient of variation of 0.25, an effect size of 0.25, an alpha of 0.05 and study power of 0.80

Sample size estimates were particularly sensitive to ES and CV parameters (Figure 4.2 a & b). The increase in number of clusters required to detect smaller effect sizes becomes progressively larger as effect size is reduced; more than twice as many clusters per study arm are needed to detect a difference of 0.20 between the malaria program and comparison program than to detect a difference of 0.30. The influence of CV on number of clusters required is relatively steady across increases in the lower range of CV values, such that for each 0.05 increase in CV between 0.15 and 0.30, approximately 5 or 6 additional units are required to maintain study power at 80%. This increases to between 8 and 10 units per 0.05 increase in CV for CV values greater than 0.30. Sample size estimates were also sensitive to baseline mortality rates, but this sensitivity was confined primarily to sample sizes below 500 and to BMRs of 40/1,000 and lower (Figure 4.2c). For smaller cluster sizes of 100 to 300 children, lower baseline mortality rates require a considerable increase in the number of required clusters.

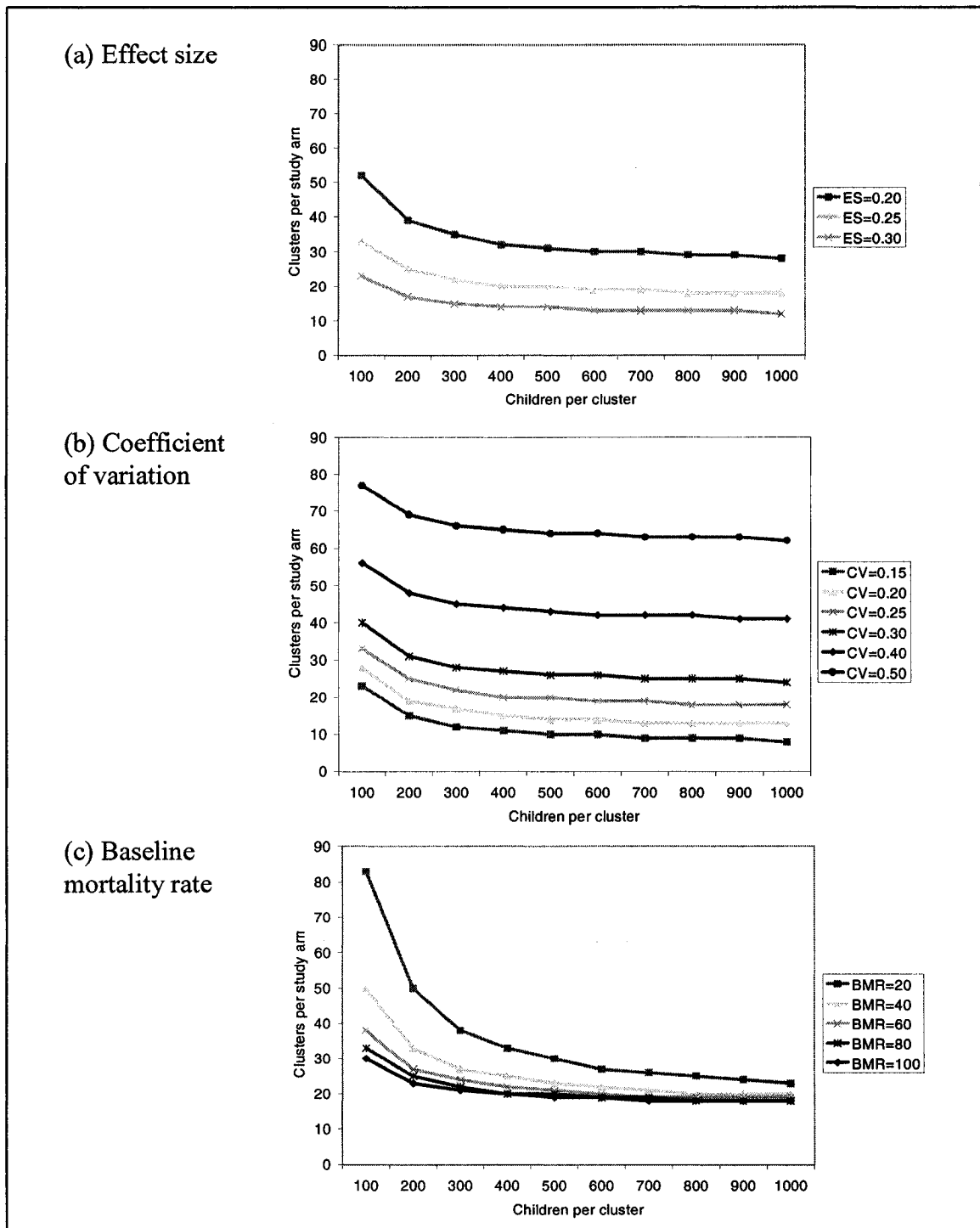


Figure 4.2 Number of clusters per arm by cluster size across selected parameters assuming a mid-range set of parameters as appropriate (a coefficient of variation of 0.25, an effect size of 0.25, a baseline mortality rate of 80/1,000 child-years), an alpha of 0.05 and study power of 0.80

As expected, increasing the number of clusters is more effective at increasing study power than selecting larger clusters (Figure 4.3). The largest gains in power through an increase in cluster size occur between 100 and 400 children, after which increasing the cluster size provides only minimal improvement. Figure 4.4 illustrates the relationship between power and between-cluster correlations for a given number of clusters. Study power is very sensitive to the estimate of between-cluster correlation and this sensitivity increases with increasing cluster size. As observed for sample size, increases in study power across cluster size for a given CV value are essentially achieved by mid-range cluster sizes.

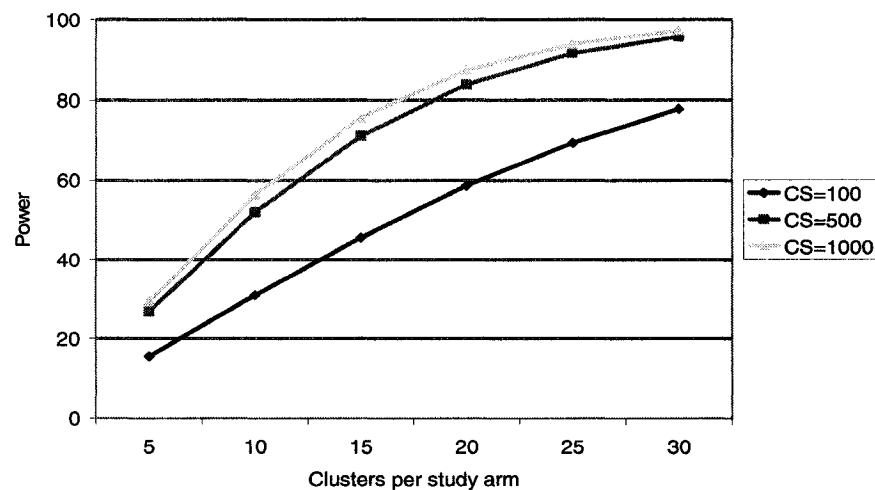


Figure 4.3 Study power by number of clusters per study arm and cluster size (CS) assuming a baseline mortality rate of 80/1,000 child-years, a coefficient of variation of 0.25, an effect size of 0.25 and an alpha of 0.05

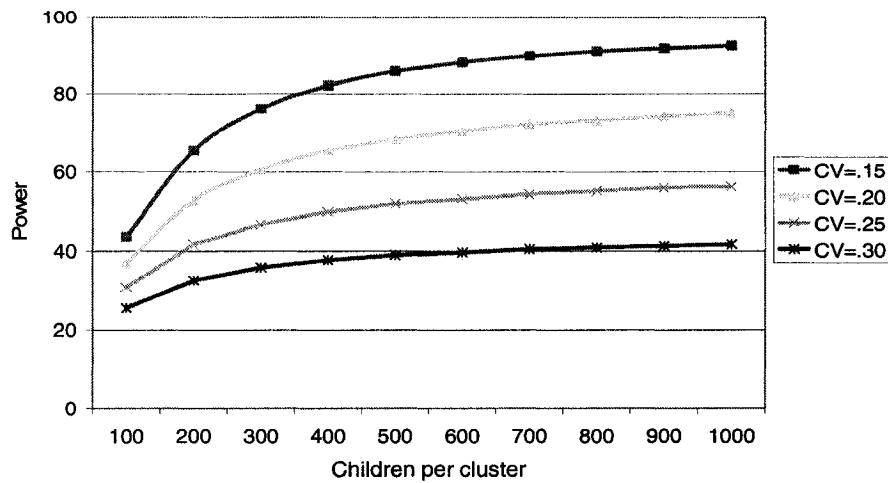


Figure 4.4 Study power by cluster size and by coefficient of variation assuming a baseline mortality rate of 80/1,000 child-years, ten clusters per study arm, an effect size of 0.25 and an alpha of 0.05

The influence of losses to follow-up of 20% and 30% on the number of required clusters is confined primarily to smaller clusters (Figure 4.5). By cluster sizes of around 500, there is little influence of losses in the range of 20% to 30% on number of required units. This parallels the decline in number of units overall as cluster sizes increase.

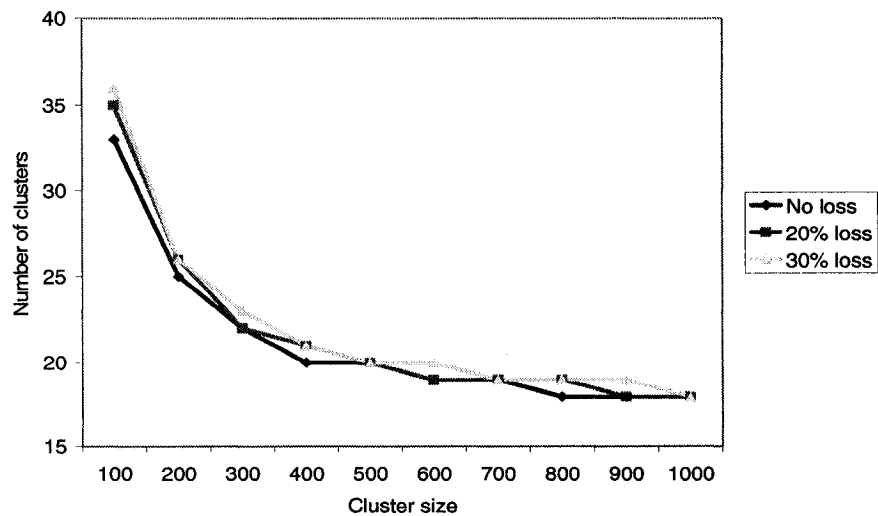


Figure 4.5 Number of clusters per study arm by cluster size and loss to follow-up assuming a coefficient of variation of 0.25, an effect size of 0.25, and an alpha of 0.05

Power declines due to losses of entire clusters were quite minimal when losses were minor. For a cluster size of 100 and enough clusters to ensure power of 80% ($n=33$; $BMR=80/1,000$; $CV=0.25$; $ES=0.25$), the loss of one cluster per arm would reduce power from 82.1% to 80.8% and three per arm to 77.9%. The same parameters with a cluster size of 500 would reduce power from 83.9% to 81.9% for one cluster lost per arm and to 77.0% for two clusters lost per arm. Larger losses disproportionately reduce power in larger clusters; for example a loss of five clusters per arm corresponds to 69.1% power for 1,000 children/cluster, to 71.2% for 500 children/cluster and to 74.9% for 100 children/cluster.

4.5 DISCUSSION

Complex community health interventions, such as the HCMMP program, are costly to develop and implement and it is important that they are evaluated in a rigorous manner¹²⁴. Cluster randomized trials are increasingly employed in the evaluation of health promotion and educational interventions targeting infectious diseases, which are difficult to deliver to individuals apart from their community context¹⁴⁶. This sub-study presents a set of proposal development guidelines for such a study based on good scientific principles along with a framework for sample size and power to guide the evaluation of the HCMMP program. This discussion will comment on the process of preparing these outputs and outline some recommendations for upcoming stages of the research process.

4.5.1 Proposal development guidelines

The HCMMP study will involve several countries following a common research protocol to answer a set of important questions, some of which require the pooling of results across study sites. At the same time, the study must deliver valuable information for local decision makers and program planners. Attempts were made to make the process of preparing the guidelines as participatory as possible to meet the expectations of researchers, international and most importantly, country level researchers and policy makers. The resulting proposal development guidelines are practical, simple, based on good scientific principles and responsive to the input of a wide range of research, programming and policy stakeholders. They also balance the somewhat conflicting aims of being both flexible to adapt to country level needs and similar enough to allow results to be combined. Although the proposal guidelines do not provide a definitive set of methods, they outline the issues that need careful consideration, lay out viable options, and identify the information required to make decisions.

The consultative meetings in Entebbe and Stockholm proved to be a successful, efficient way to obtain a diverse range of input towards designing a methodological framework addressing scientific and political angles and for building partnerships with potential research partners. The inclusion of international agencies and country level researchers and policy makers in Entebbe helped to consolidate input in a forum that allowed consensus to be reached on the major study questions and protocol elements. These were critical to establish at this early stage given their influence on other research design aspects, including sample size and the statistical analysis approaches. The continual dialogue with the WHO and with representatives from potential study sites ensured the guidelines would be suitable.

In preparing the guidelines, the aim was to be directive where necessary to allow study results to be pooled and to improve the replication and transfer of results. Several key areas were fixed including: the use of a cluster-randomized, three-arm study design, the broad intervention components, a primary outcome measure of overall under-five mortality, study duration and study phases. However, variation and flexibility in coming up with strategies and approaches was also prioritized and in many places general approaches were outlined rather than specifying parameters. Several sections of the guidelines were left open but will need to be made consistent between studies once study sites have been selected. These include the definition of care for the comparison group, the study drugs and treatment algorithm, the measurement of key secondary outcomes such as compliance and adverse events, the use of stratification, among others. The work completed to date will contribute to resolving these issues and give direction on how to proceed. It is hoped that the openness of the guidelines will promote the development of innovative approaches to several aspects of the intervention and the research process. These can then be shared for the benefit of all research teams and the study itself.

The guidelines generally followed the consensus reached at the meeting in Entebbe. However, a few notable exceptions were made during follow-up discussions in Stockholm. One key modification was the simplification of the study design. A step-wise design was proposed in the Entebbe meeting in which communities randomized to the malaria arm were to be switched to the malaria and pneumonia intervention for the second year of the study and similarly, the comparison group would be given the malaria only intervention. This was thought to complicate issues considerably and difficulties were anticipated in developing criteria to signal movement between intervention states

and in achieving the logistics (i.e. training, drug delivery mechanisms) necessary to support this change within a reasonable time period. The ethical issues that originally led to step-wise design could also be addressed in a simpler model, by having all study arms receive the intervention shown to be most effective at the end of the two year intervention phase.

A second important modification was the incorporation of the necessary formative and operational research explicitly in the protocol. Operational and evaluation objectives were separately defined and the study was divided into four phases: a formative phase; an operationalization phase; an intervention phase; and a post-intervention phase. The formative phase involves the descriptive and epidemiological studies needed to support the development of the intervention. The operationalization phase builds in the piloting and optimization of the intervention prior to its formal evaluation. Process indicators will be measured during this phase to ensure a reasonable level of coverage and acceptability has been achieved. The intervention phase is essentially the ‘evaluation’ of the study, where the optimized intervention will be allowed to run for a two year period and the primary and secondary impact indicators will be collected and analyzed during this period. The final phase, the post-intervention period, encourages the dissemination of research results to all study communities and plans for the sustainability and expansion of the ‘best care’ option. The division of the study into phases is intended to encourage country teams to identify, plan and secure funding for the formative and pilot stages of the research.

The use of the CIHR format and the CONSORT statement to direct the proposal guidelines helped ensure that the key elements of a protocol were included and that the

proposals would follow solid research principles that addressed the cluster nature of the design. However, the CIHR guidelines are intended for randomized controlled trials of therapeutic interventions, or simple interventions, not the complex, multi-component community intervention under design. As a result, they did not allow for the various phases that a community intervention study involves and that were added to the HCMMP guidelines. These preparatory and operational aspects of field trials of community based interventions are often expensive and complicated to establish. Several researchers have commented on the lack of a phased approach to the development and selection of community interventions, such as exists for therapeutic interventions (i.e. Phase I, Phase II and Phase III clinical trials)¹²⁴. Insufficient attention to initial phases can lead to the premature evaluation of programmes, producing misleading results and using considerable resources that could have been spent elsewhere. Conversely, waiting too long can result in the random assignment of groups no longer being ethical or feasible. This shortcoming of traditional trial guidelines was recently noted by the United Kingdom Medical Research Council (MRC) in their framework for the design and evaluation of complex interventions (http://www.mrc.ac.uk/pdf-mrc_cpr.pdf), as part of its series of guidelines for clinical trials. This framework suggests an iterative, phased approach consisting of five distinct stages: Preclinical (theory), Phase I (modelling), Phase II (exploratory trial), Phase III (definitive randomized controlled trial) and Phase IV (long term implementation). These phases are analogous to those for clinical trials and similar to those outlined in the HCMMP proposal guidelines.

4.5.2 Sample size framework

The importance of an adequate sample size to detect desired differences between study groups cannot be overstated, particularly for cluster trials that are generally large and expensive to conduct¹²⁴. The work presented here provides a first look at sample size requirements for the HCMMP study over a broad range of plausible situations and provides some important guidance. Based on these preliminary estimates, large numbers of children and communities will be required per site. Using estimates from a mid-range set of sample size parameters suggest that at a minimum, approximately 10,000 children will be required per site and more likely between 15,000 to 20,000, spread over more than 60 communities.

Despite efforts to search the literature and consult with country level researchers, the range of sample size parameters was wide, resulting in imprecise sample size estimates. In particular, the coefficient of variation and estimated cluster size, the two key factors which influence the magnitude of inflation of sample size in cluster samples, were not well estimated. Limited information was available on the coefficient of variation or intracluster correlation for the primary outcome of under-five mortality rates. This was not due to lack of studies measuring this outcome, as child mortality is a common outcome of child health interventions in sub-Saharan Africa, but rather to a lack of reporting of estimates or the data necessary to calculate them. In particular, four recent bednet trials conducted in Kenya¹⁴³, Ghana¹⁴⁷, Burkina Faso¹⁴⁸, and The Gambia¹⁴⁴ with a similar population group did not publish estimates. The one available estimate was instead calculated by a separate group of researchers (Hayes and Bennett) to demonstrate sample size calculations¹⁴⁰. The application of the CONSORT guidelines recently

developed for cluster-randomized trials should help ensure that future studies report this information¹³⁹.

Obtaining a good estimate of the CV is especially important for the HCMMP study since the number of children per cluster is likely to be substantial, and the impact of CV on study power increases as the cluster size increases. Available data suggests high rates of between-cluster variation for under-five mortality that were in excess of that observed for most health outcomes¹⁴⁰. In the HCMMP study, attempts should be made to obtain a more precise estimate of CV for each of the study sites and to reduce CV where possible by selecting relatively homogenous clusters with relatively stable mortality rates. This could be accomplished by collecting mortality data during the operationalization phase and would have multiple benefits for study design by providing additional information such as more precise estimates of the baseline mortality rates and estimates of covariates that influence under-five mortality at the cluster level that should be adjusted for in the analysis to improve power¹²⁴.

Due to the complex nature of the HCMMP program, obtaining a reasonable sample size in terms of number of clusters required is important. As is well known for cluster randomized trials, power and efficiency are increased by enrolling a larger number of smaller clusters. This was observed in estimates for the HCMMP study. However, cluster size is determined largely by logistical constraints and the nature of the intervention components. In the case of the HCMMP study and the countries consulted, cluster size may need to be aligned with the health delivery system. As a consequence, small clusters may not be feasible and larger units such as clusters of communities or sub-counties may be required.

Although less statistically efficient, larger clusters are associated with some advantages for the HCMMP study. As observed in the sensitivity analysis, smaller cluster sizes are considerably more vulnerable to power losses if parameters are misspecified. Larger clusters also afford a smaller number of clusters overall and are less vulnerable to losses to follow-up at the individual level. Given the uncertainty with which some of the main determinants of sample size and power are measured, a prudent approach would be to select somewhat moderate cluster sizes. Medium cluster sizes of about 300 to 500 children provide the best balance between number of clusters and number of children required. If larger clusters or higher levels of aggregation are needed for logistical reasons, the overall sample size can be minimized by sub-sampling. Sub-sampling can reduce data collection and related costs and also ensure equal cluster sizes, which has beneficial properties for statistical analysis and for sample size calculations^{149;150}.

In general, the sample size calculations were based on conservative assumptions, such as using a Bonferroni correction for multiple testing and considering high levels of between-cluster correlation. However, the sample size calculations make several assumptions that may not hold in practice and these serve to underestimate the sample size. First, they assume equal number of child-years of follow-up per cluster, which is not likely to occur. Unequal cluster sizes require further inflations of the sample size to maintain equivalent power. To account for unequal cluster sizes, all cluster sizes here should be viewed as the minimum level, to err on the conservative side¹⁴⁹. Second, these equations assume that the between-cluster variation in under-five mortality rates is equal between study arms, which may not be the case, particularly if the intervention effect is

not equal across groups. This was shown to be true in the ITN trial in The Gambia and is a likely scenario given the complexity of the HCMMP program¹⁴⁴. In determining final sample size requirements, this conservative approach will need to be balanced with potential cost implications to avoid unnecessarily increasing the cost of the study. The recruitment of additional communities is considerably more difficult and costly than recruiting additional children and may not be possible in some situations.

4.5.3 Next steps

In February 2002, immediately following the Entebbe meeting, TDR released a call for letters of interest to participate in the HCMMP study. Responses were received from seventeen research teams in twelve countries, including Burkina Faso, Ethiopia, Ghana, Kenya, Nigeria, Senegal, Tanzania, The Gambia, and Uganda. The proposal development guidelines developed here will be made available to TDR to disseminate to country teams to form the basis of their proposal submissions. Proposal submissions will be accepted by TDR until September 1, 2002 and proposals will be reviewed and recommended by members of the HCMMP working group. Selection of study sites will be made by the TDR Steering Committee on Implementation Research who will select between five to seven proposals.

Once study sites and partners are selected, a meeting with selected country teams will be organized to finalize the common research protocol. The purpose of this meeting will be to combine ideas, standardize approaches for key intervention and design issues and address outstanding issues such as comparison group composition, study drug and treatment algorithm. The sample size and power framework will be used to help finalize

study requirements for each study site. These guidelines and sample size framework make an important contribution to this ongoing design process.

5.0 CONCLUSIONS

The management of childhood fever in malarious areas with poor health infrastructure presents an enormous challenge for public health. Febrile illnesses associated with malaria and pneumonia account for a large proportion of child morbidity and mortality in sub-Saharan Africa despite the existence of safe and affordable treatments. To expand the coverage and impact of available treatments, greater attention to understanding and improving caregiver management of fever is a key priority. The three sub-studies that comprise this thesis address several issues that are central to this process.

The first sub-study demonstrated through a systematic review that caregivers perform well at detecting fever in their children using palpation. In the absence of objective measures of body temperature, caregivers can be trusted to identify most fevers. In concern for the well-being of their children, it is likely that some over diagnosis of fever by caregivers will occur. The accuracy of caregiver assessments depended on the threshold used to identify fever, and in most cases caregivers used a threshold that maximized sensitivity and specificity. These findings are encouraging for health care workers who rely on caregiver assessment in areas where several common sources of childhood fever can quickly lead to serious morbidity and mortality. Applications to a community setting and identification of other factors that influence caregiver accuracy were limited by the small number of studies overall and of studies conducted at the community level. The poor reporting of diagnostic information in primary studies confirms the need for adopting and promoting standards to improve study quality and reporting, such as those outlined in the STARD initiative.

The second sub-study identified socio-demographic and health care factors associated with caregiver fever management practices, demonstrated an important linkage between home treatment and care seeking behaviours and documented clustering of fever management behaviours within communities. Of particular interest was the differing pattern of determinants and clustering noted between home treatment and care seeking behaviours. Factors associated with home treatment with an antimalarial were individual level socio-demographic factors (caregiver education level, child age and household size). In contrast, shared variables such as distance and region were associated with care seeking behaviours. The degree of clustering within villages was greater for care seeking behaviours and could reflect a greater influence of community level factors that could not be explored in the current analysis. Complementary research is needed to better understand the sources of clustering, particularly for care seeking behaviours. These findings have important implications for the design of programs and for the evaluation of such programs. Policy makers and program planners should be cognizant of these factors when designing programs to optimize prompt and appropriate treatment of childhood fevers. Further research is also needed to explore whether these results are specific to childhood fever and their generalizability beyond the study population.

The third sub-study presented a set of proposal development guidelines and a framework for sample size estimates for a multicentre, cluster-randomized trial of an integrated approach addressing childhood fever in homes and communities. The HCMMP proposal development guidelines¹ successfully integrated input from a wide

¹ As an update, the HCMMP proposal guidelines were disseminated by TDR to country teams in more than 12 countries. In October 2002, submissions were reviewed by the TDR/IDE Steering Committee and seven proposals were selected. These research teams and the HCMMP working group will meet July 28-31 in Accra, Ghana to finalize the protocol.

variety of stakeholders and balanced the need for comparable research methodology with the need for locally relevant results. The preparation of these guidelines highlighted the need for trial guidelines for complex interventions that allow for a phased approach similar to that of clinical trials. The sample size framework revealed the vulnerability of sample size estimates and study power to imprecise input parameters, particularly among smaller cluster sizes. Collection of data during a pilot study will be important to help determine more precise estimates.

Together, the three sub-studies conducted as part of this thesis contribute to the development of home and community based programs to improve care for children with fever. This programmatic approach has the potential to extend the reach of health services to those most in need and should be carefully evaluated to ensure a strong evidence base is available to guide policy.

6.0 REFERENCE LIST

1. UNICEF. Progress since the World Summit for Children: A Statistical Review 314. 2001. New York, UNICEF.
2. Williams BG, Gouws E, Boschi-Pinto C, Bryce J, Dye C. Estimates of world-wide distribution of child deaths from acute respiratory infections. *Lancet Infect.Dis.* 2002;**2**:25-32.
3. UNICEF and World Health Organization. Africa Malaria Report 2003. WHO/CDS/MAL/2003.1093. 2003. Geneva, Switzerland, World Health Organization.
4. Roll Back Malaria. Roll Back Malaria: The Stage is Set for Impact. 2002. Geneva, Switzerland, World Health Organization. Global Threat Against Infectious Disease Threat.
5. Mulholland K. Magnitude of the problem of childhood pneumonia. *Lancet* 1999;**354**:590-2.
6. Greenwood BM, Greenwood AM, Bradley AK, Tulloch S, Hayes R, Oldfield FS. Deaths in infancy and early childhood in a well-vaccinated, rural, West African population. *Ann.Trop.Paediatr.* 1987;**7**:91-9.
7. Greenwood B. Malaria mortality and morbidity in Africa. *Bull.World Health Organ* 1999;**77**:617-8.
8. Greenwood BM, Greenwood AM, Bradley AK, Snow RW, Byass P, Hayes RJ *et al.* Comparison of two strategies for control of malaria within a primary health care programme in the Gambia. *Lancet* 1988;**1**:1121-7.
9. Department of Child and Adolescent Health and Development (CAH). Model chapter for textbooks: IMCI - Integrated Management of Childhood Illness 2001. Geneva, World Health Organization.
10. World Health Organization. The overlap in the clinical presentation and treatment of malaria and pneumonia in children: report of a meeting (Geneva, 8 April, 1991) 1992. Geneva, World Health Organization.
11. O'Dempsey TJ, McArdle TF, Laurence BE, Lamont AC, Todd JE, Greenwood BM. Overlap in the clinical features of pneumonia and malaria in African children. *Trans.R.Soc.Trop.Med Hyg.* 1993;**87**:662-5.
12. Källander K, Nsungwa-Sabiiti J, Peterson S. Symptom overlap for malaria and pneumonia- Implications for home management strategies. *Submitted* 2003.

13. UNICEF. The State of the World's Children 1996: Statistical tables. 1996. New York, UNICEF.
14. Bergevin, Y. Child Survival: Challenges and Opportunities. 2001. Presented to the Canadian Conference on International Health, November 2001, Ottawa.
15. Waldman, R., Bartlett, A., Campbell, C. C., and Steketee, R. W. Overcoming Remaining Barriers: The Pathway to Child Survival. 1996. Arlington, VA, BASICS . Current Issues in Child Survival Series.
16. Winch PJ, Leban K, Casazza L, Walker L, Percy K. An implementation framework for household and community integrated management of childhood illness. *Health Policy Plan.* 2002;**17**:345-53.
17. World Health Organization Programme for the Control of Acute Respiratory Infections. The Management of Fever in Young Children with Acute Respiratory Infections in Developing Countries. 1993. Geneva, World Health Organization.
18. Gove S. Integrated management of childhood illness by outpatient health workers: technical basis and overview. The WHO Working Group on Guidelines for Integrated Management of the Sick Child. *Bull. World Health Organ* 1997;**75** Suppl 1:7-24.
19. McCarthy PL. Fever in Infants and Children. In Mackowiak PA, ed. *Fever: Basic Mechanisms and Management*, pp 351-62. Philadelphia: Lippincott-Raven Publishers, 1997.
20. Gove S. Integrated management of childhood illness by outpatient health workers: technical basis and overview. The WHO Working Group on Guidelines for Integrated Management of the Sick Child. *Bull. World Health Organ* 1997;**75** Suppl 1:7-24.:7-24.
21. Rogier C, Commenges D, Trape JF. Evidence for an age-dependent pyrogenic threshold of Plasmodium falciparum parasitemia in highly endemic populations. *Am.J.Trop.Med.Hyg.* 1996;**54**:613-9.
22. Bouvier P, Rougemont A, Breslow N, Doumbo O, Delley V, Dicko A *et al.* Seasonality and malaria in a west African village: does high parasite density predict fever incidence? *Am.J.Epidemiol.* 1997;**145**:850-7.
23. Binka FN, Morris SS, Ross DA, Arthur P, Aryeetey ME. Patterns of malaria morbidity and mortality in children in northern Ghana. *Trans.R.Soc.Trop.Med Hyg.* 1994;**88**:381-5.
24. Velema JP, Alihonou EM, Chippaux JP, van Boxel Y, Gbedji E, Adegbin R. Malaria morbidity and mortality in children under three years of age on the coast of Benin, West Africa. *Trans.R.Soc.Trop.Med.Hyg.* 1991;**85**:430-5.

25. Smith T, Hurt N, Teuscher T, Tanner M. Is fever a good sign for clinical malaria in surveys of endemic communities? *Am.J.Trop.Med.Hyg.* 1995;**52**:306-10.
26. Greenwood BM, Bradley AK, Greenwood AM, Byass P, Jammeh K, Marsh K *et al.* Mortality and morbidity from malaria among children in a rural area of The Gambia, West Africa. *Trans.R.Soc.Trop.Med.Hyg.* 1987;**81**:478-86.
27. Nzeyimana I, Henry MC, Dossou-Yovo J, Doannio JM, Diawara L, Carnevale P. [The epidemiology of malaria in the southwestern forests of the Ivory Coast (Tai region)]. *Bull.Soc.Pathol.Exot.* 2002;**95**:89-94.
28. Bailo DA, De Serres G, Beavogui AH, Lapointe C, Viens P. Home care of malaria-infected children of less than 5 years of age in a rural area of the Republic of Guinea. *Bull.World Health Organ* 2001;**79**:28-32.
29. Katz J, West KP, Jr., LeClerq SC, Thapa MD, Khattry SK, Shrestha SR *et al.* Agreement between clinical examination and parental morbidity histories for children in Nepal. *J.Trop.Pediatr.* 1998;**44**:225-9.
30. Verhoef H, Hodgins E, West CE, Carter JY, Kok FJ. Diagnosis of fever in Africa. *Lancet* 1998;**351**:372-3.
31. World Health Organization. Report: Technical Consultation on the Management of Children with Fever in the Context of IMCI. 2001. Geneva, Switzerland, World Health Organization.
32. Smith T, Schellenberg JA, Hayes R. Attributable fraction estimates and case definitions for malaria in endemic areas. *Stat.Med.* 1994;**13**:2345-58.
33. Smith T, Schellenberg JA, Hayes R. Attributable fraction estimates and case definitions for malaria in endemic areas. *Stat.Med.* 1994;**13**:2345-58.
34. Lubanga RG, Norman S, Ewbank D, Karamagi C. Maternal diagnosis and treatment of children's fever in an endemic malaria zone of Uganda: implications for the malaria control programme. *Acta Trop.* 1997;**68**:53-64.
35. Rooth I, Bjorkman A. Fever episodes in a holoendemic malaria area of Tanzania: parasitological and clinical findings and diagnostic aspects related to malaria. *Trans.R.Soc.Trop.Med Hyg.* 1992;**86**:479-82.
36. Ng DK, Lam JC, Chow KW. Childhood fever revisited. *Hong.Kong.Med.J.* 2002;**8**:39-43.
37. Browne GJ, Currow K, Rainbow J. Practical approach to the febrile child in the emergency department. *Emerg.Med.(Fremantle.)* 2001;**13**:426-35.
38. Singhi S, Sood V. Reliability of subjective assessment of fever by mothers. *Indian Pediatr.* 1990;**27**:811-5.

39. Nwanyanwu OC, Ziba C, Redd SC, Luby SP. Palpation as a method of fever determination in Malawian children who are less than 5 years old: how reliable is it? *Ann.Trop.Med Parasitol.* 1997;**91**:359-63.
40. Banco L, Jayashekaramurthy S. The ability of mothers to read a thermometer. *Clin.Pediatr.(Phila)* 1990;**29**:343-5.
41. Schwartz N, Guendelman S, English P. Thermometer use among Mexican immigrant mothers in California. *Soc.Sci.Med* 1997;**45**:1315-23.
42. Chandramohan D, Jaffar S, Greenwood B. Use of clinical algorithms for diagnosing malaria. *Trop.Med.Int.Health* 2002;**7**:45-52.
43. Einterz EM, Bates ME. Fever in Africa: do patients know when they are hot? *Lancet* 1997;**350**:781.
44. Sirima SB, Konate A, Tiono AB, Convelbo N, Cousens S, Pagnoni F. Early treatment of childhood fevers with pre-packaged antimalarial drugs in the home reduces severe malaria morbidity in Burkina Faso. *Trop.Med.Int.Health* 2003;**8**:133-9.
45. Kidane G, Morrow RH. Teaching mothers to provide home treatment of malaria in Tigray, Ethiopia: a randomised trial. 178. *Lancet* 2000;**356**:550-5.
46. Moses LE, Shapiro D, Littenberg B. Combining independent studies of a diagnostic test into a summary ROC curve: data-analytic approaches and some additional considerations. *Stat.Med* 1993;**12**:1293-316.
47. Littenberg B, Moses LE. Estimating diagnostic accuracy from multiple conflicting reports: a new meta-analytic method. *Med Decis.Making* 1993;**13**:313-21.
48. Cochrane Methods Working Group on Systematic Review of Screening and Diagnostic Tests. Cochrane Methods Working Group on Systematic Review of Screening and Diagnostic Tests: Recommended Methods (Updated 6 June 1999). Cochrane Methods Working Group on Systematic Review of Screening and Diagnostic Tests. 1996.
49. Irwig L, Tosteson AN, Gatsonis C, Lau J, Colditz G, Chalmers TC *et al.* Guidelines for meta-analyses evaluating diagnostic tests. *Ann.Intern.Med* 1994;**120**:667-76.
50. Irwig L, Macaskill P, Glasziou P, Fahey M. Meta-analytic methods for diagnostic test accuracy. *J.Clin.Epidemiol.* 1995;**48**:119-30.
51. Midgette AS, Stukel TA, Littenberg B. A meta-analytic method for summarizing diagnostic test performances: receiver-operating-characteristic-summary point estimates. *Med Decis.Making* 1993;**13**:253-7.

52. National Health and Medical Research Council. How to review the evidence: systematic identification and review of the scientific literature. 1999. Canberra, Australia, National Health and Medical Research Council (Australia).
53. Lijmer JG, Mol BW, Heisterkamp S, Bossel GJ, Prins MH, van der Meulen JH *et al.* Empirical evidence of design-related bias in studies of diagnostic tests. *JAMA* 1999;**282**:1061-6.
54. Begg CB. Biases in the assessment of diagnostic tests. *Stat.Med* 1987;**6**:411-23.
55. Healy MJ. Statistics from the inside. 14. Diagnostic and screening tests and reference values. *Arch.Dis.Child* 1995;**72**:358-61.
56. Mitchell MD. Validation of the summary ROC for diagnostic test meta-analysis: a Monte Carlo simulation. *Acad.Radiol.* 2003;**10**:25-31.
57. Berlin JA, Laird NM, Sacks HS, Chalmers TC. A comparison of statistical methods for combining event rates from clinical trials. *Stat.Med.* 1989;**8**:141-51.
58. Jones RJ, O'Dempsey TJ, Greenwood BM. Screening for a raised rectal temperature in Africa. *Arch.Dis.Child* 1993;**69**:437-9.
59. Graneto JW, Soglin DF. Maternal screening of childhood fever by palpation. *Pediatr.Emerg.Care* 1996;**12**:183-4.
60. Hooker EA, Smith SW, Miles T, King L. Subjective assessment of fever by parents: comparison with measurement by noncontact tympanic thermometer and calibrated rectal glass mercury thermometer. *Ann.Emerg.Med* 1996;**28**:313-7.
61. Ernst TN, Philp M. Temperature assessment by parental palpation. *Am.J.Dis.Child* 1985;**139**:546-7.
62. Banco L, Veltri D. Ability of mothers to subjectively assess the presence of fever in their children. *Am.J.Dis.Child* 1984;**138**:976-8.
63. Al-Almaie SM. The reliability of mothers to assess the presence of fever in their children. *Saudi Medical Journal* 1999;**20**:696-8.
64. Alves JG, Correia JB. Ability of mothers to assess the presence of fever in their children without using a thermometer. *Trop.Doct.* 2002;**32**:145-6.
65. Whybrew K, Murray M, Morley C. Diagnosing fever by touch: observational study. *BMJ* 1998;**317**:321.
66. Boyko EJ. Re: "Meta-analysis of Pap test accuracy". *Am.J.Epidemiol.* 1996;**143**:406-7.

67. Harper R, Reeves B. Reporting of precision of estimates for diagnostic accuracy: a review. *BMJ* 1999;**318**:1322-3.
68. Suojanen JN. False false positive rates. *N.Engl.J.Med* 1999;**341**:131.
69. Bossuyt PM, Reitsma JB, Bruns DE, Gatsonis CA, Glasziou PP, Irwig LM *et al.* The STARD statement for reporting studies of diagnostic accuracy: explanation and elaboration. *Ann.Intern.Med.* 2003;**138**:W1-12.
70. Tahzib F. Fever in Africa and WHO recommendation. *Lancet* 1997;**350**:1550.
71. Nwanyanwu OC, Redd SC, Ziba C, Luby SP, Mount DL, Franco C *et al.* Validity of mother's history regarding antimalarial drug use in Malawian children under five years old. *Trans.R.Soc.Trop.Med.Hyg.* 1996;**90**:66-8.
72. Nsimba SE, Massele AY, Eriksen J, Gustafsson LL, Tomson G, Warsame M. Case management of malaria in under-fives at primary health care facilities in a Tanzanian district. *Trop.Med.Int.Health* 2002;**7**:201-9.
73. Weber MW, Mulholland EK, Jaffar S, Troedsson H, Gove S, Greenwood BM. Evaluation of an algorithm for the integrated management of childhood illness in an area with seasonal malaria in the Gambia. *Bull. World Health Organ* 1997;**75 Suppl 1**:25-32.:25-32.
74. Perkins BA, Zucker JR, Otieno J, Jafari HS, Paxton L, Redd SC *et al.* Evaluation of an algorithm for integrated management of childhood illness in an area of Kenya with high malaria transmission. *Bull. World Health Organ* 1997;**75 Suppl 1**:33-42.:33-42.
75. Verhoef H, Hodgins E, Eggelte TA, Carter JY, Lema O, West CE *et al.* Anti-malarial drug use among preschool children in an area of seasonal malaria transmission in Kenya. *Am.J.Trop.Med.Hyg.* 1999;**61**:770-5.
76. Sterne JA, Egger M, Smith GD. Systematic reviews in health care: Investigating and dealing with publication and other biases in meta-analysis. *BMJ* 2001;**323**:101-5.
77. Song F, Khan KS, Dinnes J, Sutton AJ. Asymmetric funnel plots and publication bias in meta-analyses of diagnostic accuracy. *Int.J.Epidemiol.* 2002;**31**:88-95.
78. Thera MA, D'Alessandro U, Thiero M, Ouedraogo A, Packou J, Souleymane OA *et al.* Child malaria treatment practices among mothers in the district of Yanfolila, Sikasso region, Mali. *Trop.Med.Int.Health* 2000;**5**:876-81.
79. Siebenaler ME. Taking a baby's temperature: is it common knowledge? *MCN Am.J.Matern.Child Nurs.* 1985;**10**:71.
80. Greenwood B, Mutabingwa T. Malaria in 2002. *Nature* 2002;**415**:670-2.

81. Biritwum RB, Welbeck J, Barnish G. Incidence and management of malaria in two communities of different socio-economic level, in Accra, Ghana. *Ann.Trop.Med.Parasitol.* 2000;**94**:771-8.
82. Baume C, Helitzer D, Kachur SP. Patterns of care for childhood malaria in Zambia. *Soc.Sci.Med.* 2000;**51**:1491-503.
83. Molyneux CS, Mung'Ala-Odera V, Harpham T, Snow RW. Maternal responses to childhood fevers: a comparison of rural and urban residents in coastal Kenya. *Trop.Med.Int.Health* 1999; **4**:836-45.
84. Deming MS, Gayibor A, Murphy K, Jones TS, Karsa T. Home treatment of febrile children with antimalarial drugs in Togo. *Bull.World Health Organ* 1989;**67**:695-700.
85. Hamel MJ, Odhacha A, Roberts JM, Deming MS. Malaria control in Bungoma District, Kenya: a survey of home treatment of children with fever, bednet use and attendance at antenatal clinics. *Bull.World Health Organ.*2001.;**79**.(11.):1014.-23.**79**:1014-23.
86. Nsimba SE, Warsame M, Tomson G, Masseur AY, Mbatia ZA. A household survey of source, availability, and use of antimalarials in a rural area of Tanzania. *Drug Information Journal* 1999;**33**:1025-32.
87. Nshakira N, Kristensen M, Ssali F, Whyte SR. Appropriate treatment of malaria? Use of antimalarial drugs for children's fevers in district medical units, drug shops and homes in eastern Uganda. *Trop.Med.Int.Health* 2002;**7**:309-16.
88. Amazigo UV, Brieger WR, Katabarwa M, Akogun O, Ntep M, Boatman B *et al.* The challenges of community-directed treatment with ivermectin (CDTI) within the African Programme for Onchocerciasis Control (APOC). *Ann.Trop.Med.Parasitol.* 2002;**96 Suppl 1**:S41-58.:S41-S58.
89. Uganda Bureau of Statistics and ORC Macro. Uganda Demographic and Health Survey 2000-2001. 2001. Calverton, Maryland, UBOS and ORC Macro.
90. Filmer, D. Fever and its treatment among the more and less poor in Sub-Saharan Africa. 2002. Washington, DC, Development Research Group, World Bank.
91. McCombie SC. Self-treatment for malaria: the evidence and methodological issues. *Health Policy Plan.* 2002;**17**:333-44.
92. Lindblade KA, O'Neill DB, Mathanga DP, Katungu J, Wilson ML. Treatment for clinical malaria is sought promptly during an epidemic in a highland region of Uganda. *Trop.Med.Int.Health* 2000;**5**:865-75.

93. Tarimo DS, Urassa DP, Msamanga GI. Caretakers' perceptions of clinical manifestations of childhood malaria in holo-endemic rural communities in Tanzania. *East African Medical Journal* 1998;**75**:93-6.
94. Slutsker L, Chitsulo L, Macheso A, Steketee RW. Treatment of malaria fever episodes among children in Malawi: results of a KAP survey. *Trop.Med.Parasitol.* 1994;**45**:61-4.
95. Schellenberg JA, Victora CG, Mushi A, de Savigny D, Schellenberg D, Mshinda H *et al.* Inequities among the very poor: health care for children in rural southern Tanzania. *Lancet* 2003;**361**:561-6.
96. Mosley WH, Chen LC. An analytical framework for the study of child survival in developing countries. *Population and Development Review* 1984;**10**:25-45.
97. Katz J, Zeger SL. Estimation of design effects in cluster surveys. *Ann.Epidemiol.* 1994;**4**:295-301.
98. Snow RW, Schellenberg JR, Peshu N, Forster D, Newton CR, Winstanley PA *et al.* Periodicity and space-time clustering of severe childhood malaria on the coast of Kenya. *Trans.R.Soc.Trop.Med.Hyg.* 1993;**87**:386-90.
99. Wendo C. Uganda leads the way on Africa Malaria Day. *Lancet* 2002;**359**:1494.
100. Central Intelligence Agency. The World Factbook 2002: Uganda. 2002.
101. Ministry of Health Uganda. Malaria Control and Prevention. 2002.
102. Kanya MR, Bakyaite NN, Talisuna AO, Were WM, Staedke SG. Increasing antimalarial drug resistance in Uganda and revision of the national drug policy. *Trop.Med.Int.Health* 2002;**7**:1031-41.
103. Ministry of Health Uganda. Health Infrastructure. 2002.
104. Twebaze, D. A Literature Review of Care-seeking Practices for Major Childhood Illnesses in Uganda. 2001. Arlington, Virginia, Basic Support for Institutionalizing Child Survival Project (BASICS II) for the United States Agency for International Development.
105. Bennett S, Woods T, Liyanage WM, Smith DL. A simplified general method for cluster-sample surveys of health in developing countries. *World Health Stat.Q.* 1991;**44**:98-106.
106. Hoshaw-Woodard, S. Description and comparison of the methods of cluster sampling and lot quality assurance sampling to assess immunization coverage. WHO/V&B/01.26. 2001. Geneva, Switzerland, Department of Vaccines and Biologicals, World Health Organization.

107. Ministry of Health Uganda. Health Indicators. 2002.
108. Brogan DJ. Pitfalls of using standard statistical software packages for sample survey data. In Armitage P, Colton T, eds. *Encyclopedia of Biostatistics*, John Wiley, 1998.
109. Carey VJ, Zeger SL, Diggle P. Modelling multivariate binary data with alternating logistic regressions. *Biometrika* 1993;80:517-26.
110. Bobashev GV, Anthony JC. Use of alternating logistic regression in studies of drug-use clustering. *Subst. Use. Misuse*. 2000;35:1051-73.
111. Katz J, Carey VJ, Zeger SL, Sommer A. Estimation of design effects and diarrhea clustering within households and villages. *Am.J.Epidemiol.* 1993;138:994-1006.
112. Petronis KR, Anthony JC. Perceived risk of cocaine use and experience with cocaine: do they cluster within US neighborhoods and cities? *Drug Alcohol Depend.* 2000;57:183-92.
113. Preisser JS, Koch GG. Categorical data analysis in public health. *Annu.Rev.Public Health* 1997;18:51-82.:51-82.
114. Norton EC, Bieler GS, Ennett ST, Zarkin GA. Analysis of prevention program effectiveness with clustered data using generalized estimating equations. *J.Consult Clin.Psychol.* 1996;64:919-26.
115. Tarimo DS, Minjas JN, Bygbjerg IC. Perception of chloroquine efficacy and alternative treatments for uncomplicated malaria in children in a holoendemic area of Tanzania: implications for the change of treatment policy. *Trop.Med.Int.Health* 2001;6:992-7.
116. Agyepong IA, Manderson L. The diagnosis and management of fever at household level in the Greater Accra Region, Ghana. *Acta Trop.* 1994;58:317-30.
117. Asenso-Okyere WK, Dzator JA. Household cost of seeking malaria care. A retrospective study of two districts in Ghana. *Soc.Sci.Med.* 1997;45:659-67.
118. Mugisha F, Kouyate B, Gbangou A, Sauerborn R. Examining out-of-pocket expenditure on health care in Nouna, Burkina Faso: implications for health policy. *Trop.Med.Int.Health* 2002;7:187-96.
119. Gwatkin DR. How well do health programmes reach the poor? *Lancet* 2003;361:540-1.
120. Muller I, Smith T, Mellor S, Rare L, Genton B. The effect of distance from home on attendance at a small rural health centre in Papua New Guinea. *Int.J.Epidemiol.* 1998;27:878-84.

121. Foster S. Treatment of malaria outside the formal health services. *J.Trop.Med.Hyg.* 1995;**98**:29-34.
122. Asenso-Okyere WK, Dzator JA. Household cost of seeking malaria care. A retrospective study of two districts in Ghana. *Soc.Sci.Med.* 1997;**45**:659-67.
123. Bobashev GV, Anthony JC. Clusters of marijuana use in the United States. *Am.J.Epidemiol.* 1998;**148**:1168-74.
124. Donner A, Klar N. Design and analysis of cluster randomized trials. Toronto: 2001.
125. Molyneux CS, Murira G, Masha J, Snow RW. Intra-household relations and treatment decision-making for childhood illness: a Kenyan case study. *J.Biosoc.Sci.* 2002;**34**:109-31.
126. Harrison LH, Moursi S, Guinena AH, Gadomski AM, el Ansary KS, Khallaf N *et al.* Maternal reporting of acute respiratory infection in Egypt. *Int.J.Epidemiol.* 1995;**24**:1058-63.
127. Byass P, Hanlon PW. Daily morbidity records: recall and reliability. *Int.J.Epidemiol.* 1994;**23**:757-63.
128. Hennekens C, Buring J. Epidemiology in Medicine. Boston: Little, Brown and Company, 1987.
129. Kilian A, Tindyebwa D, Gulck T, Byamukama W, Rubaale T, Kabagambe G *et al.* Attitude of women in western Uganda towards pre-packed, unit-dosed malaria treatment for children. *Tropical Medicine & International Health* 2003;**8**:431-8.
130. Mariko M. Quality of care and the demand for health services in Bamako, Mali: the specific roles of structural, process, and outcome components. *Soc.Sci.Med.* 2003;**56**:1183-96.
131. Thind A, Andersen R. Respiratory illness in the Dominican Republic: what are the predictors for health services utilization of young children? *Soc.Sci.Med.* 2003;**56**:1173-82.
132. Wears RL. Advanced statistics: statistical methods for analyzing cluster and cluster-randomized data. *Acad.Emerg.Med.* 2002;**9**:330-41.
133. Child Health and Development. Informal Consultation on Epidemiologic Estimates for Child Health, 2001. Geneva, Department of Child and Adolescent Health and Development, World Health Organization.
134. Reyes H, Perez-Cuevas R, Salmeron J, Tome P, Guiscafre H, Gutierrez G. Infant mortality due to acute respiratory infections: the influence of primary care processes. *Health Policy Plan.* 1997;**12**:214-23.

135. Tupasi TE, Velmonte MA, Sanvictores ME, Abraham L, De Leon LE, Tan SA *et al.* Determinants of morbidity and mortality due to acute respiratory infections: implications for intervention. *J.Infect.Dis.* 1988;**157**:615-23.
136. Källander K, Guenther T, Wells, G. A., Tomson, G., Bergevin, Y., and Tugwell, P. Home and community management of malaria and pneumonia - A systematic review. 2002. Submitted 2003.
137. Simpson JM, Klar N, Donner A. Accounting for cluster randomization: a review of primary prevention trials, 1990 through 1993. *Am.J.Public Health* 1995;**85**:1378-83.
138. Donner A, Brown KS, Brasher P. A methodological review of non-therapeutic intervention trials employing cluster randomization, 1979-1989. *Int.J.Epidemiol.* 1990;**19**:795-800.
139. Elbourne DR, Campbell MK. Extending the CONSORT statement to cluster randomized trials: for discussion. *Stat.Med* 2001;**20**:489-96.
140. Hayes RJ, Bennett S. Simple sample size calculation for cluster-randomized trials 165. *Int J.Epidemiol.* 1999;**28**:319-26.
141. Bellamy N, Carette S, Ford PM, Kean WF, le Riche NG, Lussier A *et al.* Osteoarthritis antirheumatic drug trials. III. Setting the delta for clinical trials--results of a consensus development (Delphi) exercise. *J.Rheumatol.* 1992;**19**:451-7.
142. Morrow RH, Smith PG. *Methods for Field Trials of Interventions Against Tropical Diseases: A Toolbox.* Baltimore: Oxford Press, 1996.
143. Nevill CG, Some ES, Mung'ala VO, Mutemi W, New L, Marsh K *et al.* Insecticide-treated bednets reduce mortality and severe morbidity from malaria among children on the Kenyan coast. *Trop.Med.Int.Health* 1996;**1**:139-46.
144. D'Alessandro U, Olaleye BO, McGuire W, Langerock P, Bennett S, Aikins MK *et al.* Mortality and morbidity from malaria in Gambian children after introduction of an impregnated bednet programme. *Lancet* 1995;**345**:479-83.
145. Sankoh OA, Ye Y, Sauerborn R, Muller O, Becher H. Clustering of childhood mortality in rural Burkina Faso. *Int.J.Epidemiol.* 2001;**30**:485-92.
146. Hayes RJ, Alexander ND, Bennett S, Cousens SN. Design and analysis issues in cluster-randomized trials of interventions against infectious diseases. *Stat.Methods Med.Res.* 2000;**9**:95-116.
147. Binka FN, Kubaje A, Adjuik M, Williams LA, Lengeler C, Maude GH *et al.* Impact of permethrin impregnated bednets on child mortality in Kassena-Nankana

- district, Ghana: a randomized controlled trial. *Trop.Med.Int.Health* 1996;1:147-54.
148. Habluetzel A, Diallo DA, Esposito F, Lamizana L, Pagnoni F, Lengeler C *et al.* Do insecticide-treated curtains reduce all-cause child mortality in Burkina Faso? *Trop.Med.Int.Health* 1997;2:855-62.
149. Yudkin PL, Moher M. Putting theory into practice: a cluster randomized trial with a small number of clusters. *Stat.Med.* 2001;20:341-9.
150. Gail MH, Mark SD, Carroll RJ, Green SB, Pee D. On design considerations and randomization-based inference for community intervention trials. *Stat.Med.* 1996;15:1069-92.

7.0 APPENDICES

APPENDIX IA: SEARCH STRATEGY

Published studies posted on databases:

Host system: Ovid Technologies

Databases: MEDLINE; HealthSTAR; Current Contents; CINAHL; Dissertation Abstracts

Latest date searched: October 28, 2002

Search terms*:

1. exp Fever/ (20,246)
2. (hyperthermia or fever or pyrexia).tw. (69,387)
3. 1 or 2 (78,709)
4. Caregivers/ (6,007)
5. (caregiver\$ or care giver\$ or parent\$ or mother\$ or father\$ or caregiver\$ or care taker\$).tw.(210,429)
6. exp Parents/ (34,367)
7. or/4-6 (226,553)
8. 3 and 7 (2,231)
9. "Sensitivity and Specificity"/ (89,174)
10. di.fs (1,075,559)
11. Predictive Value of Tests/ (44,579)
12. false negative reactions/ or false positive reactions/ (21,701)
13. (sensitivity or specificity).tw. (325,384)
14. or/9-13 (1,418,046)
15. 8 and 14 (542)
16. limit 15 to (all infant <birth to 23 months> or all child <0 to 18 years>) (380)

*Numbers in brackets indicate the number of citations from the MEDLINE search only.

Summary of search strategy results:

Database	Dates searched	Number of citations	Citations retrieved*
MEDLINE	1966-Oct. week 3 2002	380	26
HealthSTAR	1987-Sept. 2002	324	2
Current Contents	1993 week 27-week 44 2002	58	1
CINAHL	1982-Sept. week 4 2002	32	1
Dissertation Abstracts	1861- Nov. 2002	7	0

*Refers to the number of citations retrieved for MEDLINE and the number of additional citations with duplicates from preceding databases removed for the other databases.

Review of reference lists from primary articles:

Number of articles obtained through searches of reference lists of primary articles: 12

APPENDIX IB: EXCLUDED STUDIES

Summary Table:

Reason for Exclusion	Study Author & Publication Year	
Insufficient information provided to calculate sensitivity and specificity	▪ Factor et al, 2001	
	▪ Thera et al, 2000	
	▪ Rowe et al, 2000	
	▪ Bojang et al, 2000	
	▪ Dunyo et al, 2000	
	▪ Katz et al, 1998	
	▪ Lubanga et al, 1997	
	▪ Redd et al, 1996	
	▪ Smith et al, 1995	
	▪ Binka et al, 1994	
	▪ Rooth et al, 1992	
	▪ Bonadio et al, 1990	
	Caregiver assessment of fever not compared to an objective measure	▪ Talani et al, 2002
		▪ al Eissa et al, 2000
▪ Blumenthal, 1998		
▪ Bruijnzeels et al, 1998		
▪ Muhe, 1996		
▪ Watling et al, 1995		
▪ Agyepong et al, 1994		
▪ Kalter et al, 1991		
▪ McCarthy et al, 1990		
▪ Bauchner et al, 1987		
Assessors were health care workers	▪ Kilmon, 1987	
	▪ Anonymous, 1986	
Subjects assessed themselves for fever	▪ Bergeson et al, 1974	
	▪ Morgan et al, 1997	
Caregiver assessment based on volunteered reporting of fever	▪ Gomes et al, 1994	
	▪ Bialo Diallo et al, 2001	
	▪ Dunyo et al, 1997	
Time between subjective and objective measures of fever was too long	▪ Kofoed et al, 1998	

List of Excluded Studies:

Insufficient information provided to calculate sensitivity and specificity (n=12)

Factor SH, Schillinger JA, Kalter HD, Saha S, Begum H, Hossain A *et al.* Diagnosis and management of febrile children using the WHO/UNICEF guidelines for IMCI in Dhaka, Bangladesh. *Bull. World Health Organ* 2001;**79**:1096-105.

Thera MA, D'Alessandro U, Thiero M, Ouedraogo A, Packou J, Souleymane OA *et al.* Child malaria treatment practices among mothers in the district of Yanfolila, Sikasso region, Mali. *Trop. Med. Int. Health* 2000;**5**:876-81.

Rowe AK, Hamel MJ, Flanders WD, Doutizanga R, Ndoyo J, Deming MS. Predictors of correct treatment of children with fever seen at outpatient health facilities in the Central African Republic. *Am.J.Epidemiol.* 2000;**151**:1029-35.

Bojang KA, Obaro S, Morison LA, Greenwood BM. A prospective evaluation of a clinical algorithm for the diagnosis of malaria in Gambian children. *Trop.Med.Int.Health* 2000;**5**:231-6.

Dunyo SK, Afari EA, Koram KA, Ahorlu CK, Abubakar I, Nkrumah FK. Health centre versus home presumptive diagnosis of malaria in southern Ghana: implications for home-based care policy. *Trans.R.Soc.Trop.Med Hyg.* 2000;**94**:285-8.

Katz J, West KP, Jr., LeClerq SC, Thapa MD, Khatri SK, Shrestha SR *et al.* Agreement between clinical examination and parental morbidity histories for children in Nepal. *J.Trop.Pediatr.* 1998;**44**:225-9.

Lubanga RG, Norman S, Ewbank D, Karamagi C. Maternal diagnosis and treatment of children's fever in an endemic malaria zone of Uganda: implications for the malaria control programme. *Acta Trop.* 1997;**68**:53-64.

Redd SC, Kazembe PN, Luby SP, Nwanyanwu O, Hightower AW, Ziba C *et al.* Clinical algorithm for treatment of Plasmodium falciparum malaria in children. *Lancet* 1996;**347**:223-7.

Smith T, Hurt N, Teuscher T, Tanner M. Is fever a good sign for clinical malaria in surveys of endemic communities? *Am.J.Trop.Med.Hyg.* 1995;**52**:306-10.

Binka FN, Morris SS, Ross DA, Arthur P, Aryeetey ME. Patterns of malaria morbidity and mortality in children in northern Ghana. *Trans.R.Soc.Trop.Med Hyg.* 1994;**88**:381-5.

Rooth I, Bjorkman A. Fever episodes in a holoendemic malaria area of Tanzania: parasitological and clinical findings and diagnostic aspects related to malaria. *Trans.R.Soc.Trop.Med Hyg.* 1992;**86**:479-82.

Bonadio WA, Hegenbarth M, Zachariason M. Correlating reported fever in young infants with subsequent temperature patterns and rate of serious bacterial infections. *Pediatr.Infect.Dis.J.* 1990;**9**:158-60.

Caregiver assessment of fever not compared to an objective measure (n = 11)

Talani P, Samba G, Moyon G. [Management of child fever in the battle against malaria in Brazzaville] Prise en charge des fièvres de l'enfant dans le cadre de la lutte contre le paludisme-maladie a Brazzaville. *Bull.Soc.Pathol.Exot.* 2002;**95** :47-9.

al Eissa YA, al Zamil FA, al Sanie AM, al Salloum AA, al Tuwaijri HM, al Abdali NM *et al.* Home management of fever in children: rational or ritual? *Int J.Clin.Pract.* 2000;**54**:138-42.

- Blumenthal I. What parents think of fever. *Fam.Pract.* 1998;**15**:513-8.
- Bruijnzeels MA, van der Wouden JC, Foets M, Prins A, van den Heuvel WJ. Validity and accuracy of interview and diary data on children's medical utilisation in The Netherlands. *J.Epidemiol.Community Health* 1998;**52**:65-9.
- Muhe L. Mothers' perceptions of signs and symptoms of acute respiratory infections in their children and their assessment of severity in an urban community of Ethiopia. *Ann.Trop.Paediatr.* 1996;**16**:129-35.
- Watling C, Gyebi-Ofosu E, McIntyre L. Assessment of home management of fever among children in the Ashanti Region of Ghana. *J.Trop.Pediatr.* 1995;**41**:189-90.
- Agyepong IA, Manderson L. The diagnosis and management of fever at household level in the Greater Accra Region, Ghana. *Acta Trop.* 1994;**58**:317-30.
- Kalter HD, Gray RH, Black RE, Gultiano SA. Validation of the diagnosis of childhood morbidity using maternal health interviews. *Int.J.Epidemiol.* 1991;**20**:193-8.
- McCarthy PL, Sznajderman SD, Lustman-Findling K, Baron MA, Fink HD, Czarkowski N *et al.* Mothers' clinical judgment: a randomized trial of the Acute Illness Observation Scales. *J.Pediatr.* 1990;**116**:200-6.
- Bauchner H, McCarthy PL, Sznajderman SD, Baron MA, Fink HD, Forsyth B *et al.* Do mothers overestimate the seriousness of their infants' acute illnesses? *J.Dev.Behav.Pediatr.* 1987;**8**:255-9.
- Kilmon CA. Home management of children's fevers. *J.Pediatr.Nurs.* 1987;**2**:400-4.

Assessors were health care workers (n = 2)

- How good are we at guessing fever? *Ir.Med J.* 1986;**79**:15-6.
- Bergeson PS, Stienfeld HJ. How dependable is palpation as a screening method for fever? Can touch substitute for thermometer readings? *Clin.Pediatr.(Phila)* 1974;**13**:350-1.

Subjects assessed themselves for fever (n = 2)

- Morgan D, Whitworth JA, Ross A, Omoding N, Eotu H. Fever in Africa and WHO recommendation. *Lancet* 1997;**350**:1549-50.
- Gomes M, Espino FE, Abaquin J, Realon C, Salazar NP. Symptomatic identification of malaria in the home and in the primary health care clinic. *Bull.World Health Organ* 1994;**72**:383-90.

Caregiver assessment based on volunteered reporting of fever (n = 2)

Bailo DA, De Serres G, Beavogui AH, Lapointe C, Viens P. Home care of malaria-infected children of less than 5 years of age in a rural area of the Republic of Guinea. *Bull. World Health Organ* 2001;**79**:28-32.

Dunyo SK, Koram KA, Nkrumah FK. Fever in Africa and WHO recommendation. *Lancet* 1997;**350**:1550.

Time between subjective and objective measures of fever was too long (n = 1)

Kofoed PE, Dias F, Lopes F, Rombo L. Diagnosis of fever in Africa. *Lancet* 1998;**351**:373-4.

APPENDIX IIA: SELECTED QUESTIONS FROM THE COMMUNITY INTEGRATED MANAGEMENT OF CHILDHOOD ILLNESS (C-IMCI) HOUSEHOLD SURVEY

Socio-demographics:

Respondent:

- 1 = Mother of child
- 2 = Father of child
- 3 = Adult in household (15 years and above)
- 4 = Other specify _____

Age of respondent: _____ (years)

Sex:

- 1 = Male
- 2 = Female

Education level:

- 1 = None
- 2 = Didn't complete primary school (P7)
- 3 = Completed primary school (P7)
- 4 = Above primary (P7)
- 5 = Other: _____

Marital status of young child's caregiver:

- 1 = Single
- 2 = Married
- 3 = Divorced
- 4 = Separated
- 5 = Widowed

How many people live in this household? _____

Is it a boy or a girl?

- 1 = Boy
- 2 = Girl

How old is (s)he? _____ years _____ months

When was (s)he born? _____/_____/_____ (IF EXACT DATE NOT KNOWN, ENTER 15TH) DAY MONTH YEAR

Health care aspects:

Would you say that the quality of services offered at the health units in your area are good?

- 1 = Yes
- 2 = No
- 3 = Don't know

Do you get visits by health workers to your village?

- 1 = Yes
- 2 = No
- 3 = Don't know

How many visits did you receive in the village during the last six months _____
(STATE NUMBER)

Did you get visits by any health worker or community health volunteer to this household in the last 6 months?

- 1 = Yes
- 2 = No
- 9 = Forgotten/Don't know

How many times did they visit this household in the last 6 months? _____

Fever in last two weeks:

Did this child have fever in the last two weeks?

- 1 = Yes
- 2 = No
- 9 = Don't know

Fever management:

What did you do for the child at home during the illness (fever) before seeking advice or treatment outside the home?

(DO NOT READ THE ALTERNATIVES. PROBE: 'ANYTHING ELSE?' AND TRY TO CAPTURE ALL ALTERNATIVES, CIRCLE ALL MENTIONED)

- 1 = Nothing
- 2 = Performed tepid sponging
- 3 = Gave herbal medicine
- 4 = Gave Chloroquine (malariaquine, dawaquine, etc)
- 5 = Fansidar/Metakelfin
- 6 = Pain killers (Panadol/paracetamol, Aspirin)

7 = Unknown drug

8 = Other specify _____

Did you take the child YOURSELF for treatment outside the home?

1 = Yes

2 = No

After how many days or hours of fever did you go to seek care?

_____ days _____ hours

Where do you usually go to seek care?

1 = Traditional healers

2 = Private clinic

3 = Nearby government/NGO health facility

4 = Others _____

How far is it from your home up to the place where you usually go for seeking health care? _____ km

APPENDIX IIIA: SUMMARY OF ENTEBBE MEETING RESULTS

Issue	Options considered	Result
Research question	HCMMP vs. comparison OR HCMMP vs. comparison and HCMM vs. comparison	Consensus: Combined, malaria alone, and comparison
Study design	Two-arm vs. three arm; randomized vs. quasi-experimental; cluster vs. individual	Consensus: Three-arm, cluster randomized trial
Comparison group	Minimal intervention; Active intervention; Wait-list	Consensus: Minimal, two options identified
Study population	Various (6-59 months; 0-59 months, etc)	Consensus: Children 2 months to 6 years
Primary outcome	Cause-specific or overall mortality	Consensus: Overall under-five mortality
Length of follow-up	One-year vs. two-years	Consensus: Two years

APPENDIX IIIB: HCMMP PROPOSAL DEVELOPMENT GUIDELINES

Home and community management of malaria and pneumonia in children under-five: a cluster randomized controlled trial of an integrated approach

PROPOSAL DEVELOPMENT GUIDELINES

Prepared by: The UNDP/World Bank/WHO Special Programme for Research and Training in Tropical Diseases (TDR) and the Home and Community Management of Malaria and Pneumonia (HCMMP) Working Group, an interagency collaboration between researchers from the Department of Epidemiology and Community Medicine at the University of Ottawa, the Division of International Health of the Karolinska Institute, UNICEF, and the World Health Organization (WHO)

June 18, 2002

Table of Contents:

Introductory remarks	1
1. ABSTRACT	3
2. STATEMENT OF THE PROBLEM	3
3. BACKGROUND	3
3.1.1 Epidemiology of malaria and pneumonia	3
3.1.2 Overview of administrative and health care structure	3
3.1.3 Description of caretaker recognition, treatment and care-seeking behaviours for malaria and pneumonia.....	3
3.1.4 Evidence for the effectiveness of home and community management.....	3
3.1.5 Potential benefits, risks and anticipated problems and controversies regarding home and community management.....	4
3.1.6 Impact of the trial	4
4. HYPOTHESIS(ES)	4
5. STUDY OBJECTIVES	4
5.1.1 Operational objectives	4
5.1.2 Evaluation objectives	4
5.1.2.1 Primary objective	4
5.1.2.2 Secondary objectives.....	4
6. EXPERIMENTAL METHODS	5
6.1.1 Intervention and comparison overview.....	5
6.1.1.1 Study intervention and treatment protocols.....	5
6.1.1.1.1 Intervention components	5
6.1.1.1.2 Intervention parameters:.....	6
6.1.1.1.3 Roles, responsibilities and treatment algorithms	7
6.1.1.1.4 Study drugs	7
6.1.1.2 Care for comparison group	7
6.1.2 Study design	8
6.1.2.1 Rationale for a cluster design	8
6.1.2.2 Design overview	10
6.1.2.3 Study Phases.....	10
6.1.3 Target population for the intervention	11
6.1.3.1 Selection and recruitment of clusters and individuals (see sample size section and refer to chart in Appendix I)	12
6.1.3.2 Inclusion criteria.....	12
6.1.3.3 Exclusion criteria.....	13
6.1.3.4 Anticipated recruitment.....	13
6.1.4 Study outcomes	13
6.1.4.1 Primary outcome.....	13
6.1.4.2 Secondary outcomes	14
6.1.5 Data collection for outcome and process outcomes.....	15
6.1.6 Randomization.....	15
6.1.7 Blinding and protecting against bias	15
6.1.7.1 Concealment of randomization	15
6.1.7.2 Blinding to outcome indicators	16

6.1.7.3	Ascertainment bias	16
7.	STATISTICAL ANALYSIS	16
7.1.1	Sample size	16
7.1.2	Data management and analysis plan	16
8.	ETHICAL CONSIDERATIONS	17
9.	STUDY TEAM	17
10.	STUDY TIMELINE	17
11.	BUDGET AND BUDGET JUSTIFICATION	17
12.	REFERENCES.....	18
	Appendix I – Sample size determination	19

Introductory remarks

In low-income countries where malaria and pneumonia are common killers of children, where the health system and care seeking practices are constrained, investments in efforts to empower poor families in treatment and prevention of both these diseases are necessary to achieve an accelerated, profound, sustainable and equitable impact on child survival. This is especially important for sub-Saharan Africa where there has been very little overall improvement in the under-five mortality in the past decade, despite well recognized efforts.

When investigating approaches for empowering poor families in the treatment and prevention of malaria and pneumonia, several important characteristics that the two diseases share become important. Initial symptoms (fever, fast breathing, cough) and signs of severe illness (inability to eat, convulsions and difficulty breathing) are similar and indistinguishable even by trained health workers in the absence of laboratory investigations. Progression from mild symptoms to severe complications and death can be rapid and immediate, appropriate treatment is necessary to save life. Effective and inexpensive treatments exist (US\$0.05 for malaria and US\$0.30 for pneumonia) and are already being used (often inappropriately) by families for self treatment.

The UNDP/World Bank/WHO Special Programme for Research and Training in Tropical Diseases (TDR) and the Home and Community Management of Malaria and Pneumonia (HCMMP) Working Group, an interagency collaboration between researchers from the Department of Epidemiology and Community Medicine at the University of Ottawa, the Division of International Health of the Karolinska Institute, UNICEF, and the World Health Organization (WHO) invite proposal submissions from country teams to participate in multi-country trials of an integrated package of approaches for home management of malaria and pneumonia to assess operational feasibility and impact on under-five mortality.

Successful research teams will be expected to establish, support and implement an effective intervention for home and community based treatment of fever episodes presumed to be caused by malaria and pneumonia, among children living in malaria endemic communities that are under-served by health facilities and assess the impact of the intervention on under-five mortality through a three-arm, cluster randomized trial design. In addition, the teams should explore and evaluate approaches for improving and measuring adherence to antimalarial and antibiotic recommended treatment regimens, explore and evaluate appropriate ways of monitoring and reporting drug related adverse events resulting from home and community based treatment, monitor the development of drug resistance/failure in treatment response and evaluate the cost-effectiveness of the intervention.

The purpose of these guidelines is to provide potential investigators with a general outline of what to include in their proposal submission. Careful

preparation of the protocol is important to attract funding and also to establish the scientific rigor of the study. These guidelines have been developed following the outline of major granting agencies such as the Canadian Institutes for Health Research (CIHR). They are also in line with the CONSORT guidelines for randomized clinical trials that have recently been adapted for cluster trials. Researchers should also adhere to the TDR proposal submission guidelines and an application can be downloaded from the TDR website (<http://www.who.int/tdr/>).

Elaborated here are aspects that are particularly relevant for this trial and using this as a guide will facilitate approval by the TDR review committee and also aid in the search for a study sponsor. Instructions have been highlighted in italics. *Please structure your proposal using the suggested headlines.*

1. ABSTRACT

Use a structured format with the following headings: background and significance, objectives, study design, setting, study population, interventions, outcomes, sample size, interpretation of potential results, and timeline.

2. STATEMENT OF THE PROBLEM

Provide a paragraph summarizing the statement of the problem the intervention will address. Characterize the degree and extent of febrile illnesses/malaria and acute respiratory infections in children under five years and the overlap in presentation and management of these conditions, with specific reference to national and regional data as available. Include information on burden of disease measures such as incidence, morbidity, mortality and direct and indirect impacts. Describe the shortcomings associated with current treatment options for caretakers of young children with these conditions and highlight the research gaps in this area. Support for this statement can be supplied in the background section.

3. BACKGROUND

3.1.1 Epidemiology of malaria and pneumonia

Provide a brief overview of the epidemiology of malaria and pneumonia, focusing on national and regional information where possible. Include data on causes of death in children under-five, where these deaths occur, referral options, etc.

3.1.2 Overview of administrative and health care structure

Include information on population demographics, health care systems, ongoing and anticipated child survival programs, malaria and pneumonia control strategies, etc. Include both formal and non-formal sectors. Describe also available drugs, drug delivery systems, quality of drugs, drug resistance etc.

3.1.3 Description of caretaker recognition, treatment and care-seeking behaviours for malaria and pneumonia

Summarize what is known about caretaker knowledge, attitudes and practices related to common childhood illnesses from available sources (published and unpublished), focusing on locally relevant research as much as possible.

3.1.4 Evidence for the effectiveness of home and community management

Provide an overview of evidence supporting a home and community management strategy, including any information from pilot studies as applicable. Identify knowledge gaps and indicate how your study will help address them and what experience there is in addressing these.

3.1.5 Potential benefits, risks and anticipated problems and controversies regarding home and community management

Identify and discuss any potential direct or indirect adverse events associated with an integrated approach and indicate how the study will monitor and minimize these.

3.1.6 Impact of the trial

Explain how the results of this study will impact on child health policies and program delivery.

4. HYPOTHESIS(ES)

At an individual level, early treatment for malaria and pneumonia episodes with efficacious and appropriate treatment taken properly provides clinical benefit and averts death. The question is whether providing a community level program can translate into a significant child survival benefit.

Elaborate and expand on the above theme to develop a clear hypothesis.

5. STUDY OBJECTIVES

The following operational and evaluation objectives have been suggested; expand upon these in your proposal as necessary. Note that the secondary objectives listed are suggestions, not requirements; consider these and others when identifying your secondary objectives.

5.1.1 Operational objectives

- To develop and implement an intervention for home and community management of malaria and pneumonia that is feasible, acceptable, achieves high coverage and adherence, and is safe
- To implement a community based vital statistics monitoring system in the study communities

5.1.2 Evaluation objectives

5.1.2.1 Primary objective

- To evaluate the impact on under-five mortality achievable through a home and community approach for presumptive treatment and management of malaria and pneumonia

5.1.2.2 Secondary objectives

- To evaluate the impact on incidence of severe malaria and pneumonia
- To assess the impact on severe anemia
- To measure the cost-effectiveness of the intervention
- To assess the extent to which this intervention reaches the very poor and most at risk

To achieve these objectives, a three-arm, cluster randomized study is envisioned with a primary outcome of under-five mortality. You will be asked to describe the care for the intervention and comparison groups in section 6.1.1 and the composition of the three-arms and the unit of randomization will be defined in section 6.1.2. In later sections, you will be asked to define the target population, study outcomes, randomization and blinding procedures and outline your statistical analysis plan.

6. EXPERIMENTAL METHODS

6.1.1 Intervention and comparison overview

In this section, you will first be providing details about the study intervention groups (section 6.1.1.1) and secondly the care for the comparison group (section 6.1.1.2).

6.1.1.1 Study intervention and treatment protocols

An integrated approach focuses on simultaneously addressing the factors at the home and community level that contribute to child mortality from malaria and pneumonia. Thus, the following should be considered as goals when developing the components of the intervention:

- Well informed communities within which caretakers of children under-five have appropriate knowledge and skills*;
- Appropriate efficacious drugs close to every home where an under-five lives;
- A continuous and uninterrupted supply of drugs within the community through trained well supervised residents (mothers, CHWs, retail shop owners, volunteers etc) and at the local health centre;
- A referral awareness/system for the very ill and those with complications;
- A supply, support supervision and re-training programme; and
- A mechanism for reaching the most vulnerable and those hard to reach

*The desirable behavioural modifications from the intervention that will influence the outcome and impact most are: early suspicion of illness onset (fever, cough and fast breathing); early care seeking, prompt administration of fully efficacious drugs; complete adherence to treatment dosing and schedule; and early identification of severe illness, complications or lack of improvement and prompt referral (assuming a functioning and accessible health service).

6.1.1.1.1 Intervention components

Provide a detailed description of how you will design an intervention that will achieve these behavioural changes through a multi-faceted, home and community-level programme with four main components outlined below. Be sure to include any formative studies that are necessary. Provide details on how

these will be implemented in the different arms of your study and the duration and timing of each component.

- Information, Education, Communication (IEC)
 - *Describe message content and evidence base for its development*
 - *Describe any outstanding gaps in knowledge regarding factors that influence caretaker treatment and care-seeking behaviours for febrile and ARI illnesses in children and indicate how these will be addressed to ensure the messages are appropriate*
- Training of and skills support for study drug distributors and caregivers
 - *Indicate who will be trained and who will be responsible for their training and supervision*
 - *Specify involvement of formal and informal health providers, paying particular attention to sources where caretakers commonly obtain drugs and care from, such as drug sellers, traditional healers etc.*
 - *Describe training approach (i.e. group training versus skill impacting techniques enhancement strategies, person-to-person communication etc)*
 - *Describe required training materials and how they will be developed or adapted from existing materials*
 - *Outline the timing, duration and frequency of training for health workers and for those responsible for educating mothers*
- An effective and sustainable study drug supply system
 - *Identify who will supply caretakers with study drugs, how they will obtain them and what mechanism will be put in place to maintain a constant supply of study drugs, record keeping, drug records*
- An integrated referral and monitoring system
 - *Provide an overview of existing health services and referral systems in proposed study areas (both formal and informal)*
 - *Describe linkages between the proposed home and community management programme and existing health services (both formal and informal)*

6.1.1.1.2 Intervention parameters:

The definition of parameters is important to identify when the intervention has reached a level of implementation adequate to be considered 'optimized'. The following are suggested parameters for reaching optimization; please consider these and any others you feel appropriate:

- At least X% of febrile illness episodes receiving study treatment within 24 hours of symptom onset. *Provide an estimate of the percentage of febrile episodes that could be treated within 24 hours within an*

optimized program. Use experiences with pilot projects or other research to support your estimate if available.

- The treatment should consist of drug(s) with at least 85% clinical efficacy (see section 6.1.1.1.4 on study drugs).
- Adherence to full course of treatment of at least X% for antibiotics and X% for antimalarials. *Check with your country level IMCI implementers regarding the validity of these proposed levels of adherence and suggest appropriate options that can be achieved in your study.*

6.1.1.1.3 Roles, responsibilities and treatment algorithms

Traditionally treatment algorithms have been developed for those who will provide treatment. For this study, action guidelines and algorithms for additional levels, including the home, are needed. *Indicate the different roles for the home, the community, the health worker and the health centre. At the level of the health worker, there are IMCI and other community recommendations, explain how you will adapt them for your study.*

6.1.1.1.4 Study drugs

Suggest options of potential study drugs in your country. Follow national treatment guidelines as much as possible if you consider them to be efficacious drugs. Be sure to provide convincing efficacy data to support your selections. Include also estimates of levels of resistance, acceptability, side effects, possible drug interactions etc. and describe packaging, labelling and quality assurance measures.

6.1.1.2 Care for comparison group

The decision regarding what should be considered a minimal level of care for the comparison group raises both ethical and study design issues. Comparing the home and community management of malaria and pneumonia against 'usual care' will provide information on the true impact in communities that normally cannot be reached. However, there are currently efficacious strategies available and it may be unethical to deny comparison villages access to these. Ways to balance these considerations include:

- provide something the same to the intervention and comparison group so that its effect will balance over the study groups;
- reinforce an existing program with known benefits but that is unlikely to impact on the primary outcome (under-five mortality); or
- leave the comparison group as it is and accept that at the end of the study, the comparison group will receive the intervention if it is shown to be effective and acceptable.

An in-depth discussion of these issues took place at the recent workshop on an integrated approach to home and community management in Entebbe and

promoting bednets and providing the health facility in the study area with basic drugs were measures that emerged as two possible options.

Consider these options and others and describe what and how you would provide care and services to the comparison group in an appreciable, but non-interfering manner.

1. **BEDNETS:** The meeting in Entebbe proposed that creating awareness and demand for bednets in the intervention and comparison arms and ensuring that bednets are available for purchase in all study communities would have some advantages, including:

- the comparison group receives a known effective intervention;
- standardization in the use of bednets may be obtained across the study groups (important since bednet use is the largest potential source of confounding);
- delivery of bednets would provide lessons to inform control of malaria in the country;
- evidence on the efficacy and effectiveness of bednets is available to help adjust sample size estimates appropriately; and
- ethical arguments for providing bednets to the comparison group would be addressed.

Potential limitations of this approach include difficult logistics of promoting bednet use and the likely increase in sample size requirements.

2. **HEALTH FACILITY DRUG SUPPLY:** The meeting also proposed that an option to ensure a continuous supply of efficacious antimalarials and antibiotics for the treatment of pneumonia in children at the health facility level be considered. No additional training would be provided to the staff or mothers in this regard and drugs would not be pre-packaged. Since the premise is that the intervention will improve timeliness and appropriateness of treatment, this option is advantageous as it will help to ensure that inadequate drug supply at the health centre in the comparison group is not a confounder.

6.1.2 Study design

6.1.2.1 Rationale for a cluster design

The gold standard for evaluating the impact of an intervention or treatment is the individually randomized controlled study. Deviations from the gold standard almost always result in a loss of study efficiency and decreased ability to limit

bias and maximize precision. Therefore, it is important to outline the rationale for choosing a design other than the gold standard.¹

Consultations with research experts and local researchers point towards the use of a cluster design as the most feasible approach, based on the scientific, practical and ethical reasons outlined below:

- Cluster action of the intervention. Although the effects of treatment will be seen on an individual level, it is reasonable to expect that improved treatment and clearance of infection could affect the transmission of disease and therefore act at the level of the community as well.
- Treatment cluster contamination. Contamination of the control group is of particular concern for interventions that aim to change social behaviour or impart knowledge. A key component of the intervention is to educate mothers/caretakers to recognize signs and symptoms of malaria and pneumonia and provide prompt and appropriate treatment. It is likely that if such a program were implemented on an individual basis, there would be considerable opportunity for mothers in the control group to be exposed to the intervention via contact with other mothers.
- Enhance subject adherence. It is plausible that adherence, an integral part of the proposed intervention, may be enhanced by interaction between mothers and community members (social pressure, expected behaviour)
- Administrative convenience. Randomizing at a community level will allow concentration of personnel (i.e. trainers of mothers and/or community health workers)
- Political. Consent from community leaders as well as local politicians and perhaps national governments will be required prior to participation of individual community members and will be easier to explain and obtain in this design.
- Access to data. Estimates of vital statistics and other important data are more likely to be available at a cluster rather than an individual level.
- Additional data on effectiveness under real life situations. In most cases, educational or prevention and treatment programs are implemented at the community level or broader level.
- You may add other justifications and concerns that are applicable in your study setting.

Please discuss in choosing a cluster design which reasons are applicable and describe. If you feel an individually randomized trial represents the best approach in your setting, please indicate so, addressing the above concerns.

¹ Murray, D.M. (1998). Design and Analysis of Group Randomized Trials. Oxford University Press, New York; Hutton, J.L. (2001). Are distinctive ethical principles required for cluster randomized controlled trials? *Statistics in Medicine*, 20:473-488

6.1.2.2 Design overview

Three main principles should guide the identification of an appropriate design to evaluate an integrated approach: 1) a randomized design is required to obtain the best protection against bias; 2) an appropriate comparison group must be included; and 3) the design should address comparison versus malaria and pneumonia together.

One design option for consideration is a three-arm design in which clusters are randomized to one of three arms (see Figure 1):

- malaria and pneumonia treatment (MP treatment group 1);
- malaria only treatment (M treatment group 2); and
- comparison group.

Such a design would provide evidence regarding the effectiveness of malaria home and community management programmes as well as allow an assessment of an approach addressing malaria and pneumonia together.

As the direct evaluation of the added benefit of jointly addressing malaria and pneumonia would require a sample size that may not be feasible to obtain within a single study site, analysis for this comparison will only be possible on a multi-centre level, pooling results from all sites together.

6.1.2.3 Study Phases

It is expected that reaching the study objectives will take place over several phases as outlined below:

Formative phase: Prior to arm-specific activities, descriptive, formative and epidemiological research will likely be necessary to support the intervention development as well as activities in the comparison arm. *Suggest some specific studies and activities that you feel will be important at this stage, bearing in mind the two specific objectives (operational and evaluative) and the components of the intervention.*

Operationalization phase: *If you decide to propose a three-arm study, provide specifics regarding activities to operationalize the intervention according to arm to achieve optimal implementation.* Intervention parameters were described above (see section 6.1.1.1.2).

Intervention phase: Once intervention arms have reached optimization, the study will allow two years of full implementation. It is anticipated that during these two years, at least one full malaria transmission cycle will be covered at all sites. At the end of the two-year period, the impact on mortality and other outcomes will be analyzed, disseminated and published.

Provide details of activities that will be ongoing in the different arms of your study during this phase.

Post-intervention phase: If either intervention shows a significant impact on mortality, a process will be initiated to provide this intervention in the control arm and to integrate it into the health delivery system of the country. *Indicate how this process would be achieved in your setting.*

In your study timeline (see section 10), indicate the time you expect each phase will require.

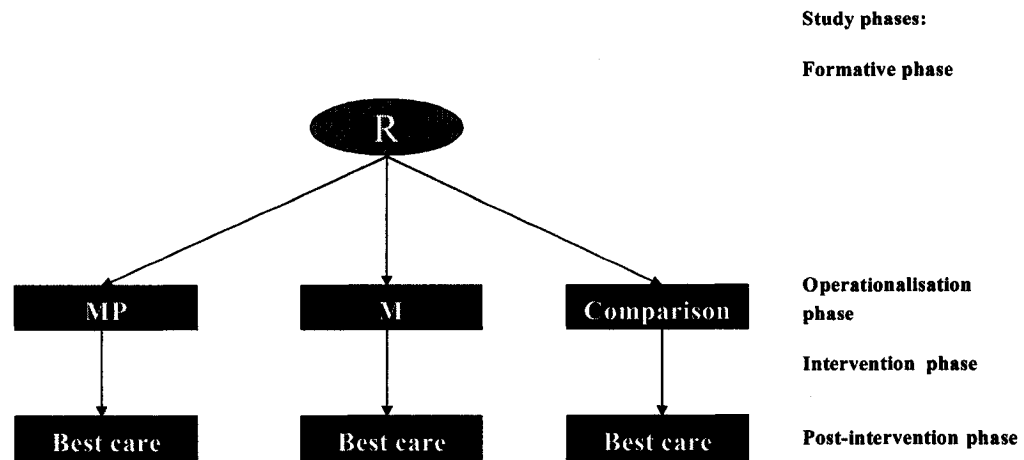


Figure 1. Overview of study design and phases

6.1.3 Target population for the intervention

This intervention is targeted at under-five children who have poor access to adequate health care services and treatment because of financial and/or economic constraints or because they live too far away from the health services. These children suffer a high rate of mortality due to malaria and pneumonia and are usually resident in rural areas.

Plasmodium falciparum has a rapid disease progression and in areas where this infection occurs all year round there is substantial incidence of malaria and high death rates in under-fives who do not receive prompt appropriate treatment. In these same settings, pneumonia incidence is high, and the two diseases together cause the majority of under-five deaths. Poor access to diagnostic facilities has resulted in presumptive treatment as the case management policy for both conditions, and therefore this makes it possible to treat outside the formal health services.

Expand upon these and other descriptions as they apply to your national and local situation.

6.1.3.1 Selection and recruitment of clusters and individuals (see sample size section and refer to chart in Appendix I)

Include the following information in this section:

- Identification of a feasible unit of randomization (i.e. village, village clusters, sub-district, ward etc) and a description of its administrative structure, population size, estimated number of children under five years, linkages to policy or other relevant structures for health care delivery etc. Provide a rationale for selecting this unit level (i.e. why not a smaller unit). Keep in mind that a smaller unit size will usually lead to a smaller overall sample size (see Appendix I for additional information).
- Selection strategy for identifying potentially eligible clusters indicating the procedure (random sample versus purposeful selection) and practical issues of consideration (i.e. natural separation between clusters to avoid contamination, ongoing and planned programs that could impact on primary outcome, available data on important factors related to eligibility criteria such as mortality and morbidity, treatment and care-seeking behaviour in relation to febrile and ARI illnesses, etc)
- Identify:
 - Who will be contacted to request involvement and obtain informed consent (i.e. unit head, caretaker, household head etc).
 - When contact will be made and by whom
 - Procedure for soliciting participation
- Timing of informed consent (see randomization section 6.1.6)

6.1.3.2 Inclusion criteria

Itemize and provide a rationale and operational definition for inclusion criteria under each of the following levels: unit of randomization and child.

The following are some suggested criteria; *please add to these as appropriate.*

Unit of randomization:

- Located in an area where malaria transmission is high and stable (*provide estimates in support of these levels of transmission*)
- *Plasmodium falciparum* is the predominant strain of malaria
- Under-five mortality rate is high (i.e. 100 per 1,000 child-years)
- Illnesses consistent with malaria and/or pneumonia constitute a considerable proportion of under-five mortality and morbidity (i.e. 30% or more)

Child:

- Between 2 –59 months at start of sustenance phase, or born and enter study area aged two months

6.1.3.3 Exclusion criteria

As above. Suggested exclusion criteria include:

Unit of randomization:

- Malaria and pneumonia are not perceived to be priority issues
- Urban areas (*provide definition, i.e. areas located along road network, in big trade centres etc.*)
- Security issues or previous history of poor participation in community studies

Child:

- Very young infants (i.e. aged 0 to 2 months)

6.1.3.4 Anticipated recruitment

Indicate expected participation rate for both randomized clusters and individuals and support your estimates with experience from other similar trials as available. Describe any initiatives planned to maximize and maintain recruitment of clusters and individuals (i.e. control clusters offered the intervention program if it is shown to be successful etc) and describe any planned back-up strategies should recruitment fail to meet expectations. Provide an estimate of how much time and the type of resources will be needed to reach your required recruitment. How will you promote your program to sustain the interest of the community, of the community workers, of the mothers? How will you popularize the intervention? Discuss compensation and remuneration issues and how they will impact on the study. Discuss also how you will make your program competitive with other approaches, monitor adverse effects and counter possible detractions they could cause to the program.

6.1.4 Study outcomes

6.1.4.1 Primary outcome

The primary outcome measure will be the difference in mortality rates in children aged 2 to 59 months between each intervention arm and the comparison arm. The reason for excluding those aged below 2 months is because it is not expected that this intervention to have an impact on deaths in this neonatal and early infant period. *Using this as your primary outcome, indicate the minimal important difference and how you will ascertain mortality outcomes.*

6.1.4.2 Secondary outcomes

1. Mortality preceded by an acute febrile illness.

Ideally, this would be the best measure of the impact of the intervention as it focuses on the symptoms we are expecting to target. However, it involves verbal autopsies which are notoriously difficult to conduct on a large scale. *If you are in a position to measure this febrile related mortality with realistic additional logistical support, describe how this will be done.*

2. Incidence of severe illness (severe fever, not able to drink or breast-feed, convulsions, abnormally sleepy or difficult to wake, chest indrawing, other danger signs)

This incidence should be monitored at the health facility as well as the community level. *Describe how you will measure this indicator.*

3. Severe anemia

Severe anemia is a common serious complication associated with malaria. *Describe how you will monitor severe anemia in your study. Consider both the health facility and the community.*

4. Cost-effectiveness of the interventions

The public health relevance of this intervention will have to be weighed against other interventions in relation to costs and effectiveness. Information on the cost of the intervention will help guide decision makers, implementers, and donor agencies. It is strongly encouraged that this data be collected. *If you will collect it, describe your methodology.*

5. Equity analysis

Very often interventions do not reach the very poor and those on the margins of society. In this study, it would be desirable to demonstrate that this is a pro-poor strategy. It would therefore be important to assess the extent to which this intervention reaches this population. *If you are in a position to collect this information, describe your methodology.*

6. Adherence to treatment and management regimens

Adherence is a key determinant of the effectiveness of this intervention and it needs to be enhanced and monitored in both qualitative and quantitative ways. *Describe how you will enhance adherence in your study as well as how you will define and monitor it.*

7. Adverse events

This intervention involves widespread availability and use of antimalarials and antibiotics, which may have rare side effects which need to be monitored. *Indicate how you will communicate this possible risk to communities, how you will minimize the risk, and how you will monitor these adverse events.*

8. Drug resistance

There is always the fear that increasing availability of drugs, especially if rational use cannot be guaranteed, that drug pressure will increase as will risk of development of resistance. It would be desirable to monitor development of resistance to study drugs. *It may be possible to gather this data from existing sources but if you are going to conduct specific drug monitoring in your study sites, describe your methods.*

6.1.5 Data collection for outcome and process outcomes

Provide a summary of the primary and secondary response variables scheduled for collection during baseline and follow-up surveys. Outline briefly how you will enumerate village residents and indicate whether you will utilize key informants at the village level to supplement information from village censuses. Indicate who will be responsible for data collection and how they will be trained. Also indicate how data for process indicators will be collected.

6.1.6 Randomization

Include the following in this section:

- Describe the methods that you will use to generate the random allocation sequence (i.e. simple random allocation versus restricted random allocation)
- Indicate whether matching or stratification will be used. If so, detail your rationale and indicate what variables clusters will be matched or stratified on.
- This intervention has community, household and individual elements. For each of these levels, and for the intervention and evaluative activities, indicate who will provide consent, how and when it will be sought and confirmed. This is a very important ethical issue and details are important.
- There has been experience to show that randomization becomes more acceptable if communities are involved in the process. For example, calling a meeting of leaders for the assignment to intervention and comparison arms helps to promote understanding and acceptability of the outcome. You may wish to consider using a similar approach.
- Describe how you will ensure that the intensity of the intervention is equal in all the intervention area at all times.

6.1.7 Blinding and protecting against bias

6.1.7.1 Concealment of randomization

Describe the mechanism planned to implement the random allocation procedure to ensure that eligible clusters are randomized in an unbiased manner. Outline how you will ensure separation between personnel responsible for generating the assignment schedule and those executing the assignment.

6.1.7.2 Blinding to outcome indicators

The Data Monitoring Committee (DMC) should not be aware from which study arm the reported outcomes (e.g. deaths, anemia rates etc.) come from. *Describe the methods you will use to keep the DMC blind in this manner. Where blinding is not possible, explain why and provide details on alternative options or implications of not blinding on interpretation of study results.*

At the operational level, in order to avoid bias, it is necessary to keep separate those implementing the program and those evaluating the program. *Describe how you will ensure this separation.*

6.1.7.3 Ascertainment bias

Please outline measures planned to minimize bias in the ascertainment of primary and secondary outcomes.

7. STATISTICAL ANALYSIS

7.1.1 Sample size

In order to demonstrate an impact on mortality, it is obvious that a large number of children in the target group will need to be included in the study. The number of children you will need to follow-up will be influenced by many factors such as the under-five mortality in the study areas, the size of the cluster you select, the magnitude of the difference you wish to measure, in and out migration, adherence to the intervention and evaluation procedures and the level of significance you plan to achieve.

Using statistical formulae that account for cluster randomization, supported by specific assumptions and health delivery realities relevant to your setting, and taking the factors above, propose a sample size for the target group that you will include in your study for the three arms and the number of clusters you will include. You might wish also to propose sample sizes for the secondary outcomes you intend to measure. Show the formulae, assumptions and the steps you have taken to arrive at this sample size.

A chart of sample size ranges and other background information has been prepared to aid in sample size determination if required (see Appendix I).

7.1.2 Data management and analysis plan

The different phases of the study are outlined in section 6.1.2.4; each of these phases involves different studies and will generate datasets and outcome indicators. *Provide a comprehensive data analysis plan with the statistical analysis that you will perform for your intended process and outcome measures,*

including any sub-group analysis. If you plan to conduct cost-effectiveness, equity or other analyses, describe your analysis plan for these as well.

Describe how you will deal with losses to follow-up and in and out-migration in your analysis and also those who met the eligibility criteria but could not participate (e.g. if you were using a sulfa drug and had to exclude children with allergies etc).

8. ETHICAL CONSIDERATIONS

Outline the process for ethical approval and describe arrangements to maintain confidentiality where relevant. Identify any additional ethical concerns and describe how these will be addressed.

9. STUDY TEAM

Core study team should include the following expertise and experience:

The leadership of this project should be provided by a principal investigator and a co-principal investigator, one from a university or research institution and the other from the Ministry of Health. Other members of the investigating team should include a pediatrician, health promotion expert, social scientist and an epidemiologist/statistician. Involving the national drug authority at an early stage is recommended.

Keep in mind that those involved in developing and implementing the intervention should be different from those evaluating it.

CVs and evidence of consent to participate in the study should be provided for all investigators.

Describe how you will recruit, train and supervise the research team necessary to fulfill your objectives. Identify short and long term research capacity strengthening needs especially those in social science for which designated support is a possibility.

10. STUDY TIMELINE

Provide a GHANT chart for your study, especially giving details for the first 12 months.

11. BUDGET AND BUDGET JUSTIFICATION

If your proposal is successful, and funding is available, funds will be dispensed in 12-month allotments. It is therefore important that you budget on a year-to-year basis. *Provide a detailed itemized, justified, budget for the first year, as well as an overall budget for the entire study. Follow the TDR budget form subheadings and use locally appropriate estimates which you can justify.*

12. REFERENCES

Provide a list of supporting references.

Below are some references that you might find helpful to read.

Donner, A. & Klar, N. Design and analysis of cluster randomization trials in health research. Hodder Headline Group, London: 2000.

Hayes, R.J., Alexander, N.D.E., Bennett, S. & Cousens, S.N. Design and analysis issues in cluster-randomized trials of interventions against infectious diseases. *Statistical Methods in Medical Research*, 2000; 9. 95-116.

Kirkwood, B.R., Cousens, S.N., Victoria, C.G., de Zoysa, I. Issues in the design and interpretation of studies to evaluate the impact of community-based interventions. *Trop. Med. Int. Health*, 1997, 2 (11), 1022-1029.

WHO. The overlap in the clinical presentation and treatment of malaria and pneumonia in children: report of a meeting. 1993. Geneva, World Health Organization.

Rasmussen Z, Pio A, Enarson P. Case management of childhood pneumonia in developing countries: recent relevant research and current initiatives. *Int J. Tuberc. Lung Dis.* 2000;4:807-26.

Mtango, F.D.E., Neuvians, D. Acute respiratory infections in children under five years. Control project in Bagamoyo District, Tanzania. *Trans. R. Soc. Trop. Med. Hyg.*, 1986;80, 851-858.

Sazawal, S., Black, R.E. Meta-analysis of intervention trials on case management of pneumonia in community settings. *Lancet*, 1992; 340, 528-533.

Kidane G., Morrow RH. Teaching mothers to provide home treatment of malaria in Tigray, Ethiopia: a randomised trial. *Lancet* 2000;356:550-5.

Pagnoni, F., Convelbo, N., Tiendrebeogo, J., Cousens, S., Esposito, F., 1997. A community-based programme to provide prompt and adequate treatment of presumptive malaria in children. *Trans. R. Soc. Trop. Med. Hyg.*, 91, 512-517.

Marsh, V.M., Mutemi, W.M., Haaland, A., Watkins, W.M., Otieno, G., Marsh, K. Changing home treatment of childhood fevers by training shop keepers in rural Kenya. *Trop. Med. Int. Health*, 1999, 4, 383-389.

Dunyo, S.K., Afari, E.A., Koram, K.A., Ahorlu, C.K., Abubakar, I., Nkrumah, F.K. Health centre versus home presumptive diagnosis of malaria in southern Ghana: implications for home-based care policy. *Trans. R. Soc. Trop. Med. Hyg.*, 2000, 94, 285-288

Diallo, A.B., De Serres, G., Béavogui, A.H., Lapointe, C., Viens, P. Home care of malaria-infected children of less than 5 years of age in a rural area of the Republic of Guinea. *Bull. World Health Organ*, 2001; 79, 28-32.

Appendix I – Sample size determination

Table 1 may be used to get a rough estimate of what sample size (both number of clusters and total number of children) you may be looking at given the number of children aged 2-59 months per cluster you feel is most appropriate and the baseline mortality rate among this age group you expect. These estimates were calculated assuming a significance level of 0.05, a power of 80%, a coefficient of variation of 0.25, a mortality reduction of 25% compared to the comparison group for the malaria arm, and a study duration of two years.

Notes on calculations:

The formula used to determine sample size to compare mortality rates under conditions of cluster sampling was that suggested by Donner and Klar²:

$$k = \frac{(Z_{\alpha/2} + Z_{\beta})^2 (\lambda_1 + \lambda_2)}{t(\lambda_1 - \lambda_2)^2} \times IF_t$$

where:

$$IF_t = 1 + \frac{CV^2(\lambda_1^2 + \lambda_2^2)t}{\lambda_1 + \lambda_2}$$

and:

k is the number of clusters required per treatment group;
α is the significance level (α = 0.05; Z_{α/2} = 2.24);
β is desired power (β = 0.80; Z_β = 0.84);
λ₁ is the average mortality rate for the intervention group;
λ₂ is the average mortality rate for the control group;
t is child-years of observation in each cluster; and
CV (coefficient of variation) is an estimate of the intrinsic variation in mortality rates between clusters.
IF is the inflation factor (or design effect)

Higher CV values indicate greater between-cluster variability and lead to a larger inflation of the required sample size. The formula used to calculate k is for a completely randomized design; however if matching or stratification are employed, the reported estimates will be conservative since these strategies act to reduce CV.

To account for the additional comparison made using the three arm design, a Bonferroni correction has been applied to the alpha level (0.05/2=0.025). The

² Donner, A. & Klar, N. Design and analysis of cluster randomization trials in health research. Hodder Headline Group, London: 2000.

total number of clusters for a two arm design is 3k. The impact of a programme addressing malaria alone would be expected to be less than that for a programme addressing both conditions. Therefore, for the three arm design, the total number of clusters is derived by taking the number of clusters computed for malaria alone. This ensures there will be enough power to make the malaria only comparison.

Table 1. Sample size estimates in total number of clusters (C) and total number of children (N) by baseline mortality rate and number of children aged 2-59 months per cluster

Children 2-59 months per cluster	Baseline child mortality rate per 1,000 child-years									
	20		40		60		80		100	
	C	N	C	N	C	N	C	N	C	N
100	249	24,900	150	15,000	114	11,400	99	9,900	90	9,000
200	150	30,000	99	19,800	81	16,200	75	15,000	69	13,800
300	114	34,200	81	24,300	72	21,600	66	19,800	63	18,900
400	99	39,600	75	30,000	66	26,400	60	24,000	60	24,000
500	90	45,000	69	34,500	63	31,500	60	30,000	57	28,500
600	81	48,600	66	39,600	60	36,000	57	34,200	57	34,200
700	78	54,600	63	44,100	57	39,900	57	39,900	54	37,800
800	75	60,000	60	48,000	57	45,600	54	43,200	54	43,200
900	72	64,800	60	54,000	57	51,300	54	48,600	54	48,600
1,000	69	69,000	60	60,000	57	57,000	54	54,000	54	54,000

Notes: C = total number of clusters required in study; N = total number of children required in study; Mortality rates are those for children aged 2-59 months and expressed per 1,000 child-years