Optimising methodology for the elicitation of participant-reported data relating to drug safety in resource poor settings

Elizabeth Allen

Thesis presented in fulfilment of the requirements for the degree Doctor of Philosophy in the Division of Clinical Pharmacology, Faculty of Health Sciences, University of Cape Town

December 2015

Supervisor: **Professor Karen I Barnes, University of Cape Town**

Co-supervisors: **Dr Clare IR Chandler, London School of Hygiene and Tropical Medicine, London, UK**

Dr Lynn Atuyambe, Makerere University, Kampala, Uganda

Optimising methodology for the elicitation of participant-reported data relating to drug safety in resource poor settings

I hereby declare that the work on which this thesis is based is my original work (except where acknowledgements indicate otherwise) and that neither the whole work nor any part of it has been, is being, or is to be submitted for another degree in this or any other university. In the case of multi-authored published or unpublished papers I was the lead author and my contribution as agreed to by the relevant co-authors is outlined in each publication. In addition I have detailed contributions in the preface to this thesis and in the introduction to each chapter.

I empower the university to reproduce for the purpose of research either the whole or any portion of the contents in any manner whatsoever, citing publication references where relevant.

Signed by candidate

Elizabeth Allen
December 2015

ABSTRACT

In addition to treating symptomatic patients, malaria prevention and elimination requires giving antimalarial drugs to asymptomatic or uninfected individuals. This shifts the harm:benefit balance and heightens the importance of accurately defining drug safety. Large data sets, including those pooled from multiple sources, are needed to understand uncommon adverse drug reactions. Interpreting individual studies, comparisons between studies and pooled datasets can be compromised, however, by inadequate or varied methods of safety data collection. Specifically, questioning methods may influence participants' reports of medical history, adverse events (AEs) and non-study medications.

A Cochrane systematic review synthesised literature on research comparing methods for eliciting AEs from trial participants. A global online survey investigated how antimalarial researchers collect these data, and mixed-methods were used to identify barriers to accurate and complete reporting in South African and Tanzanian antimalarial-antiretroviral drug interaction trials. Focus group discussions were conducted in Ghana, Kenya and Uganda with women in a drugs exposure pregnancy registry to examine barriers to reporting at antenatal clinics, and how they might be overcome.

The review included thirty-three studies in various therapeutic areas showing that more specific questioning increases the number of AEs reported by trial participants. Survey responses of 52 antimalarial researchers in 25 countries evidenced a range of methods to obtain AEs, medical histories and non-study drug reports. Qualitative data revealed that the trial context is influential and that detailed questioning facilitated participants' recognition and consideration of what to report. Non-reporting is due to forgetting, not knowing drug names, considering which information is relevant or significant to themselves or trial/healthcare workers, the potential consequences of reporting, and perceiving verbal responses inferior to what blood test results can show. Pregnant women's improved relationship with antenatal staff facilitated information-sharing and registry tools helped overcome problems with recall and naming of medicines.

This project provides evidence of the substantial impact of different questioning methods on safety assessments. The results should contribute to developing a framework for researchers when planning globally-relevant, yet context-specific, antimalarial drug safety data collection strategies, and enhance efforts to pool data from multiple sources.

PREFACE

As per provision 6.7 in the General Rules for the degree of Doctor of Philosophy (PhD) of the University of Cape Town, and as approved in 2010 by the University Doctoral Degrees Board on submission of my proposal with prior approval of my supervisor, this thesis includes four published papers and one publication-ready paper as chapters:

- Allen EN, Chandler CIR, Mandimika N, Barnes KI. Eliciting adverse effects data from participants in clinical trials (protocol). The Cochrane Library 2013;9.2. This protocol was peer reviewed and published by the Cochrane Collaboration in 2013. The final review conducted under the published protocol, presented in **Chapter 3**, will be submitted to the Cochrane Collaboration for peer review in 2015.
- Allen EN, Chandler CIR, Mandimika N, Pace C, Mehta U, Barnes KI. Evaluating harm associated with anti-malarial drugs: a survey of methods used by clinical researchers to elicit, assess and record participant-reported adverse events and related data. Malaria Journal 2013;12:325. **Chapter 4**.
- Allen EN, Mushi AK, Massawe IS, Vestergaard LS, Lemnge M, Staedke SG, Mehta U, Barnes KIB, Chandler CIR. How experiences become data: The process of eliciting adverse event, medical history and concomitant medication reports in antimalarial and antiretroviral interaction trials. BMC Medical Research Methodology. 2013;13(1):140. **Chapter 5**.
- Allen EN, Gomes M, Yevoo L, Egesah O, Clerk C, Byamugisha J, Mbonye A, Were E, Mehta U, Atuyambe LM. Influences on participant reporting in the World Health Organisation drugs exposure pregnancy registry; a qualitative study. BMC Health Services Research. 2014;14(1):525. **Chapter 6**.

As for many clinical research projects, particularly those in public health, I worked within multi-disciplinary teams. I have therefore described my contribution at the start of each chapter. For all papers I was the lead and corresponding author; under supervision I conceived the research, conducted (or supervised the conduct of) the research, managed and analysed the data outputs and wrote the first draft of the manuscript. I then sent the draft manuscripts to the co-authors for their review, and made the submissions to the relevant journal. I also drafted responses to reviewer comments, managed the required input from the co-authors and re-submitted subsequent versions to the journals. My supervisor confirms that all aspects of the thesis overwhelmingly reflect my own work.

ACKNOWLEDGEMENTS

This thesis is concerned with understanding more about the conversations we have with research participants about their health and medicine use, and considering if there is room for improvement, and harmonisation, of these interactions within or between studies. The topic by its nature meant talking with participants, which was the greatest privilege; when feeling overwhelmed by how to think about what I was trying to achieve, I often went back to my own interview recordings to listen to their experiences and challenges (within and outside of the research) and, most of all, their palpable enthusiasm for being involved in this project. My main thanks are therefore to these participants. My other participants were clinical researchers. I asked some searching questions about a topic that most probably have in the back of their minds but haven't thought about in detail. I would like to thank them for their time, honesty and commitment to further work.

I would sincerely like to thank my supervisor, Karen Barnes, for her mentorship over many years; her gentle but always perceptive encouragement, and for allowing me to pursue these research interests. It takes courage to delve into what is essentially the accuracy of one's own data, but she never saw this as a roadblock, believing instead that we must do whatever we can to find better ways to produce high quality research. I would like to express huge thanks to Clare Chandler from The London School of Hygiene and Tropical Medicine, for fundamentally changing the way I think about supposedly quantitative research. She brought an invaluable blend of epidemiology and anthropology, critical for successful public health research. I would also like to thank Lynn Atuyambe from Makerere University, who I was privileged to work with on the study presented in chapter 6, for his insights (the fruits of his own valuable research) into the often difficult world of being a pregnant woman in Africa.

I have much appreciation for my colleagues at UCT, the Cochrane Collaboration and the World Health Organisation (WHO), particularly Ushma Mehta and Melba Gomez for allowing me to nest my research questions into their Pregnancy Registry, the funders (the ACT Consortium though a grant from the Bill and Melinda Gates Foundation to the LSHTM, the WHO and my own department). Finally, thank you to my family for allowing me to travel extensively at times and for putting up with my preoccupation with this degree.

Finally, I would like to thank the examiners for their thoughtful and useful comments which have undoubtedly strengthened the thesis.

ABBREVIATIONS

ACT	artemisinin-based combination therapy
ADR	adverse drug reaction
AE	adverse event
AIDS	acquired immune deficiency syndrome
AL	artemether-lumefantrine
ANC	antenatal clinic
ANNSERS	Antipsychotic Non-Neurological Side-Effects Rating Scale
ARV	antiretroviral
ASEX	Arizona sexual experience scale
B	blank form-type enquiry (see glossary for how this was used in chapter 3)
BARS	Barnes Akathisia Rating Scale
CCOAT	Collaborating Centre for the Optimisation of Antimalarial Therapy
CIOMS	Council for International Organizations of Medical Sciences
CL	checklist-type enquiry (see glossary for how this was used in chapter 3)
CNS	central nervous system
COMET	Core Outcome Measures in Effectiveness Trials
CONSORT	Consolidated Standards of Reporting Trials
CTC	Common Toxicity Criteria
CTCAE	Common Terminology Criteria for Adverse Events
D	diary-type enquiry (see glossary for how this was used in chapter 3)
EMA	European Medicines Agency
FDA	Food and Drug Administration
FGD	focus group discussion
GCP	Good Clinical Practice
HIV	human immunodeficiency virus
HMG CoA	3-hydroxy-3-methylglutaryl-coenzyme A

ICH	International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use
IDI	in-depth interview
IMP	investigational medicinal product
INT	interview-type enquiry (see glossary for how this was used in chapter 3)
LUNSERS	Liverpool University Neuroleptic Side-effect Rating Scale
MCC	Medicines Control Council
MIP	Malaria in Pregnancy (Consortium)
MRC	Medical Research Council
NME	new molecular entity
O	open-type enquiry (see glossary for how this was used in chapter 3)
OMERACT	Outcome measures in rheumatology
OTC	over-the-counter
PK	pharmacokinetic
PRO	patient-reported outcome
R	rating scale-type enquiry (see glossary for how this was used in chapter 3)
RCT	randomised controlled trial
SAE	serious adverse event
SAFTEE	Systematic Assessment For Treatment Emergent Events
TB	tuberculosis
UCT	University of Cape Town
UKU-SERS	Udvalg for Kliniske Undersogelser (UKU) side-effect rating scale
WHO-TDR	World Health Organisation- the Special Programme for Research and Training in Tropical Diseases
VAS	visual analogue scale
WWARN	WorldWide Antimalarial Resistance Network

GLOSSARY

Adverse event (AE)	Any untoward or unfavourable medical or psychological occurrence in a participant, including any abnormal laboratory finding, symptom or disease. An AE does not necessarily have a causal relationship with the research or any risk associated with the research (1).
Adverse drug reaction (ADR) (largely considered synonymous with adverse effect)	In the pre-approval clinical experience with a new medicinal product or its new usages, particularly as the therapeutic dose(s) may not be established: all noxious and unintended responses to a medicinal product related to any dose should be considered adverse drug reactions. The phrase responses to a medicinal product means that a causal relationship between a medicinal product and an adverse event is at least a reasonable possibility, i.e. the relationship cannot be ruled out. Regarding marketed medicinal products: a response to a drug which is noxious and unintended and which occurs at doses normally used in man for prophylaxis, diagnosis, or therapy of diseases or for modification of physiological function (1).
Blank form-type enquiry (B)	A term I used in the systematic review (Chapter 3) to describe where participants were given a blank piece of paper, or equivalent, to indicate which AEs they had experienced.
Check list-type enquiry (CL)	A term I used in the systematic review (Chapter 3) to describe where participants were given a pick-list, or questionnaire, of potential AEs to indicate their health or medicine experience from.
Clinical research	For this thesis, synonymous with health research as defined by the South African National Health Act 61 of 2003: any research which contributes to knowledge of: the biological, clinical, psychological or social processes in human beings; improved methods for the provision of health services; human pathology; the causes of disease; the effects of the environment on the human body; the development or new application of pharmaceuticals, medicines and related substances; and the development of new applications of health technology (2).
Clinical trial/study	Any investigation in human subjects intended to discover or verify the clinical, pharmacological and/or other pharmacodynamics effects of an investigational product(s), and/or to identify any adverse reactions to an investigational product(s), and/or to study

	absorption, distribution, metabolism, and excretion of an investigational product(s) with the object of ascertaining its safety and/or efficacy. The terms clinical trial and clinical study are synonymous (1).
Concomitant medication	See non-study medicine.
Diary-type enquiry (D)	A term I used in the systematic review (Chapter 3) to describe where participants were given some kind of a diary to complete with the AEs they experienced.
Delphi method	A method used to obtain the most reliable opinion through consensus of a group of experts by subjecting them to a series of questionnaires interspersed with controlled opinion feedback (3).
Efficacy	The ability of a medicine or medical technology to bring about the intended beneficial effect on individuals in a defined population with a given medical problem, under ideal conditions of use (4).
Effectiveness	A measure of the effect of a medicine (or medical technology) is purported, or is represented, to have under conditions for the use prescribed, recommended or labelled (4).
Exposure	A substance that a study participant, particularly a pregnant woman taking part in a pregnancy registry, may have used (i.e. a medicine, herb, alcohol etc.)
Good clinical practice (GCP)	A standard for the design, conduct, performance, monitoring, auditing, recording, analyses, and reporting of clinical trials that provides assurance that the data and reported results are credible and accurate, and that the rights, integrity, and confidentiality of trial subjects are protected (1).
Interview-type enquiry (INT)	A term I used in the systematic review (Chapter 3) to describe where participants were interviewed to elicit AEs experienced.
Investigational product	A pharmaceutical form of an active ingredient or placebo being tested or used as a reference in a clinical trial, including a product with a marketing authorization when used or assembled (formulated or packaged) in a way different from the approved form, or when used for an unapproved indication, or when used to gain further information about an approved use (1).
Harm	The totality of possible adverse consequences of an intervention or

	therapy; they are the direct opposite of benefits, against which they must be compared (5).
Non-study medication	A medicine used by a participant in a study other than that being investigated or otherwise used as prescribed in the study protocol.
Observational study	A study where the investigator does not control the therapy, but observes and evaluates the results of ongoing medical care (6).
Patient-reported outcome (PRO)	Any report of the status of a patient's health condition that comes directly from the patient, without interpretation of the patient's response by a clinician or anyone else (7).
Pharmacoepidemiology	The study of the use and the effects of drugs in large numbers of people (6).
Pharmacovigilance	The science and activities relating to the detection, assessment, understanding and prevention of adverse effects or any other drug related problem (8).
Pregnancy registry	Observational studies measuring the association between drugs taken during pregnancy and the pregnancy outcome (9).
Previous or prior medication	Medicines taken by a study participant before they are enrolled into a study.
Prospective study	A study performed simultaneously with the events under study; namely, outcomes have not yet occurred as of the outset of the study (6).
Randomised controlled trial (RCT)	A study where the investigator randomly assigns patients to different therapies/study arms (6).
Rating scale-type enquiry (R)	A term I used in the systematic review (Chapter 3) to describe where participants were given some kind of a tool to rate their AE (e.g. a visual analogue scale).
Resource-poor setting	Synonymous with low and middle-income (LMIC) according to the World Bank classification (10)
Risk	The probability that something will happen (6).
Safety	Substantive evidence of an absence of harm. (5).
Serious adverse event (SAE)	Any untoward medical occurrence that at any dose, results in death, is life-threatening, requires inpatient hospitalization or prolongation of existing hospitalization, results in persistent or

	significant disability/incapacity, or is a congenital anomaly/birth defect (1).
Tolerability	The degree to which overt adverse effects can be tolerated by the subject (11).
Unexpected adverse drug reaction	An adverse reaction, the nature or severity of which is not consistent with the applicable product information (e.g. Investigator's Brochure for an unapproved investigational product or package insert/summary of product characteristics for an approved product) (1).
Validity	The degree to which an assessment (e.g. questionnaire or other instrument) measures what it purports to measure (6).

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

:

LIST OF SUPPLEMENTARY TABLES AND FILES

CHAPTER 1: BACKGROUND

This thesis will examine factors influencing the participant-reported data that are integral to the assessment of drug safety in resource-poor settings. In this chapter I will introduce the concepts of safety and harm, and the fields of pharmacovigilance and pharmacoepidemiology and their particular importance for resource-poor settings. I will then describe two specific study designs that are relevant to the thesis and how participant-reported data contribute to the measurement of safety outcomes. This will be followed in Chapter 2 with a review of literature pertinent for the studies conducted as part of the thesis, and justification for their rationales and objectives.

1.1 Pharmacovigilance and pharmacoepidemiology

Pharmacovigilance is “the science and activities relating to the detection, assessment, understanding and prevention of adverse effects or any other drug related problem” (8). The scope of this relatively new field is now accepted to extend to herbs, traditional and complementary medicines, blood products, biologics, medical devices and vaccines. Moreover, its remit embraces lack of efficacy, abuse or misuse of medicines, substandard medicines and interactions (12-14). Pharmacovigilance does not need to be prescriptive or bound by a particular method but rather informed by the stakeholders’ goals of identifying and preventing harm in a particular context (15). Nonetheless, drug manufacturers must comply with certain rigorous standard requirements in order to apply to regulatory authorities for registration, demonstrating a satisfactory safety profile alongside efficacy and quality. While the term 'safety' is commonly used to infer the absence of harm, no drug is without harmful effects; participants in clinical studies are monitored for the presence or absence of harm and the evaluation of the probability for harm informs the risk associated with the drug (4). Any such harms must be in an acceptable balance with its benefits, cognizant of the context of use, which will include the intended indication, population, exposure and cost. Each of these factors are considered from the various stakeholders' perspectives, whether the regulatory authorities, policy-makers, prescribers or consumers; it is thus a relative concept (16).

Bringing a new molecular entity (NME) to the market is a long, complex, and therefore inherently costly endeavour. Planning and managing the necessary clinical trials demands

continual choices about the most appropriate objectives and designs. These enable timely, scientifically-valid decisions concerning whether to continue or terminate development. An early decision to halt due to findings relating to safety, tolerability, efficacy, or feasibility will save costs and prevent further unnecessary animal and human exposure. However, this resolution must be founded on relevant and accurate data, or future patients may be deprived of a valuable treatment option. A typical clinical development plan encompasses small exploratory phase I trials (in healthy volunteers or patients depending on the NME's properties) to determine important initial safety, tolerability, pharmacokinetic (PK) profiles and, if possible, pharmacodynamic effects (17). These are followed by Phase II trials to establish the suitable dose(s) and preliminary efficacy in about 100 to 500 patients with the target condition, and finally large pivotal trials in about 1000 to 10000 patients to confirm efficacy. It is obligatory to accumulate evidence about the drug's characteristics relating to safety throughout.

It is accepted that, even when well designed and conducted, pre-registration clinical trials will not detect all adverse drug reactions (ADRs) as the numbers of participants are too small to detect rare effects, the timeline may be too short for detecting delayed or latent effects, and eligibility criteria are too strict to reflect real-world use. For instance, investigators may design eligibility criteria for a trial to minimise potential harm based on what is known about the drug thus far, although similar restrictions may not then be applied once the drug is adopted in routine practice after registration. These deficiencies pose considerable challenges to regulatory authorities negotiating the delicate balance between harms and benefits, in order to recommend drugs that are optimal for individual patients and populations alike (18). Therefore, once a drug is registered, onus is put on manufacturers to continue pharmacovigilance, employing pharmacoepidemiological methods to assess safety and tolerability of the drug in larger numbers of people (19). Passive methods, such as spontaneous reporting by providers or consumers and the interrogation of healthcare data bases, are used to detect safety signals, which are an excess of previously unidentified adverse effects that are suspected to be drug-related. Active methods, such as further trials, registries or drug event monitoring programmes, investigate the safety of long-term and real-world use of a drug.

1.2 The importance of pharmacoepidemiology for resource-poor regions

Aside from the manufacturer's own studies, others are conducted by various interested parties, such as non-governmental or public health organisations and academia. This research often fills important gaps in the evidence relating to the use of the drug by those not included in the pre-registration trials, including off-label use; for example by pregnant women, children or the elderly, those with various co-morbidities and taking concomitant medications not already investigated, or merely those in geographical areas not relevant to the manufacturer's marketing plans (20, 21) . While some such trials contribute to regulatory dossiers, others focus on informing clinicians and policymakers about the optimal use of the registered drug in the intended target populations. Although the situation is changing, particularly with the advent of product development partnerships between industry and non-profit organisations, pharmaceutical companies are less likely to invest in resource poor regions (22). Meanwhile, drugs must, of course, be researched there, in the populations that carry the highest burden of the disease under study. Therefore, while previously these regions may have been excluded from trials or targeted for trials only for economic reasons, they may be the setting of choice for particular studies. There will be some overlap between well-resourced and poorly-resourced settings in regard to factors influencing the collection of safety data, but may be more difficult to conduct due to a range of issues: weak regulatory oversight, poor infrastructure, fewer experienced trial sites, more challenging follow-up of participants and cultural practices and expectations, such as differing beliefs about medicines (23, 24). The concomitant use of other medicines, including traditional therapies, may also be more extensive, though under-documented (25, 26). Post-registration pharmacovigilance programmes are also often less developed, thus regulatory authorities may be less capable of monitoring safety effectively once the drug is distributed (15, 27). Consequently there is the potential for safety concerns to remain undetected before widespread damage.

1.3 The specific case of malaria

Despite significant expansion of malaria intervention coverage this disease still accounts for a large burden of morbidity and mortality in resource-poor regions; this is particularly so for Africa whose population suffers 90% of global malaria-related deaths, the majority being children (28). Antimalarial drugs are pivotal components of malaria control, both in terms of treatment and prevention of the disease. The current first line treatment for

uncomplicated malaria is an artemisinin-based drug in combination with a long-acting partner, a so-called artemisinin-based combination therapy (ACT) (29). The artemisinins as a class are considered relatively safe (30). However there are still questions to be answered, such as the long term safety of repeated doses in children, their safety in pregnant women who are particularly at risk from malaria and may inadvertently be exposed to contra-indication drugs, and the safety of healthy 'at risk' populations given preventive treatment or exposed during elimination strategies (29-32). In addition, due to recent evidence of resistance by Plasmodium *falciparum* parasites to the artemisinins, there is a constant need to keep a healthy pipeline of new drugs under development (33). Thus, aside from clinical trials, various post-registration pharmacoepidemiology studies and other pharmacovigilance surveillance activities are required to adequately monitor safety (34-36). Each such activity has strengths and weaknesses and the choice as to the most appropriate activity will rely on issues such as the specific study objectives, available funding, infrastructure and other issues that affect the feasibility of implementation. For instance, the increasing use of healthcare databases and cohort-event monitoring for detecting signals of potential harm are less likely to be possible in malaria-endemic settings due to financial and technological constraints (15). While spontaneous reporting mechanisms may be relatively less costly that post-marketing clinical trials, they suffer from under- or incomplete reporting (37). This thesis will focus on two types of such studies; clinical trials and prospective pregnancy registries. This is largely pragmatic in that these activities are the focus of the candidate's workplace, and which highlighted the need for this research.

1.4 Study designs relevant for the thesis

1.4.1 Clinical trials

Clinical trial design is dependent on the phase of drug development, the drug's properties, the particular research question, and sometimes practical or ethical considerations (38). Trials may be randomised, double-blind, placebo or active controlled, open label, single or parallel group, cross-over, factorial or otherwise. Sometimes the drug under investigation is the only intervention, or there may be a combination of drugs used to explore potentially beneficial or harmful combinations or interactions. Participants may also range from healthy volunteers to those with complex or life-threatening medical conditions, and could be of different ages, gender, race, or cultures. Despite this variety, data about potential

drug-related harms are expected to be collected in all trials, although the safety outcome(s) are more often secondary than primary objectives (39). Listings of AEs and potential ADRs relating to registration of a drug will ultimately be tabulated in the safety section of a regulatory dossier, using both investigator and manufacturer terms, such that corresponding rates by exposure may be described or calculated (39).

1.4.2 Pregnancy registries

Pregnancy registries are observational cohort studies designed to measure the association between exposure(s) (whether a drug or otherwise) and the pregnancy outcome, and sometimes the development of the child (9, 40-42). The aim may therefore be assessment of adverse effects on the health of the pregnant women herself, and/or the foetus, neonate or child. This will impact on the length of follow-up required. While the scope for design is not as wide as for clinical trials, registries may be specific to one or more exposure(s) or even enrol all pregnant women regardless of exposure at the study start. A key characteristic, however, is enrolment of women before the outcomes are known in order to prevent bias in the assessment of the association with the exposure(s) (43).

1.5 Serious ADRs impacting on morbidity and mortality

A blurring of boundaries between pre- and post-registration pharmacovigilance has led to earlier and more comprehensive risk management by a wide variety of experts involved in bringing a drug to the market and beyond (19, 44) . Theoretically, if communication between all those involved is maintained, it should be feasible to make timely decisions about effects on an exposed population. As data mass, confidence in the drug builds, or signals of harm are detected for further investigation. In reality, even in developed regions where regulatory compliance is likely to be the best, serious drug-related harm has been detrimental to public health and resulted in product withdrawals or restrictions (45). Major revisions in the regulatory framework have been made in response to specific incidences of severe harm, e.g. phocomelia with thalidomide in the 1960s, oculomucutaneous syndrome with practolol in the 1970s, arrhythmias with terfenadine in the 1990s, myocardial infarction and stroke with rofecoxib in the 2000s (6). However, it is not always clear if data leading to product withdrawals could have been detected earlier by the use of better clinical research methodologies (4). Sometimes the reasons are multifactorial. A cytokine storm triggered by TeGenero's TGN1412, that caused life-threatening and permanently

disabling ADRs in six healthy volunteers, was partially due to a lack of interpreting pre-clinical data according to already established recommendations (46). However, the case also highlighted a need for a thorough review of how first-in-human trials are conceived, designed, conducted, monitored and analysed to prevent a similar catastrophe. This was because of the increase in the numbers of trials of investigational medicinal products (IMPs) with inherently greater potential for harm, such as biological agents with novel mechanisms of actions and IMPs with highly species-specific activity or that target the immune system (47). Thus advances in medicine present challenges in knowing how to prevent or monitor adverse drug effects using traditional approaches.

1.6 Less serious ADRs tipping the harm:benefit pendulum and influencing adherence

While serious ADRs have an evident negative impact on morbidity and mortality, other seemingly minor drug-related problems can also have a significant undesirable impact on individuals, but which is often less measurable, attributable or even valued. Even relatively ill patients may find it hard to tolerate what clinicians consider as mild ADRs if they are persistent, so minor clinical issues can therefore cause distress and even affect adherence (48, 49). The assessments of harm versus benefit, and their risk, are not simple, despite the development of mathematical models for the latter, as these concepts may be considered and valued differently because of differing perspectives, knowledge or context (50, 51). Non-adherence due to ADRs may also be contingent on how a drug is used; a minor ADR in a patient may be unacceptable if the drug is introduced as a disease preventive strategy in a healthy population, such as during mass drug administrations (52, 53), when potentially more people with either contra-indicated or understudied conditions will be exposed (54). This may negatively impact adherence and undermine important public health interventions. It is therefore important to also pay attention to these less severe ADRs during drug development and post-registration, or public health programmes may be compromised.

1.7 Building blocks underpinning the assessment of harm

All clinical trials and pharmacoepidemiology studies rely on the collection of untoward medical occurrences that are not necessarily caused by the study drug. These adverse events (AEs) are assessed either on a case-by-case basis or by aggregate statistical synthesis

in order to provide evidence of which are likely to be ADRs (4, 55). AEs are not, however, evaluated in isolation as the participant's exposure to non-study medications, their underlying medical history and concurrent conditions, are all important pieces of information needed for judging whether an AE is likely to be caused by the investigational drug or not. These variables are also, of course, essential components for confirming study eligibility criteria and interpreting efficacy outcomes, where relevant.

As there are some limitations of current methods for assessing drug safety, it is important to question how the required data are ascertained. While there are various processes involved, this thesis focuses on examining one particular aspect, indeed one of the first steps in identifying relevant data: participant reports about their past medical histories, changes in their health after exposure to a study drug, and use of other medicines. These variables together make for an assessment of drug-related harm may sometimes be measured objectively, such as through a laboratory or other test, medical examination or a drug chart entry, or they may be a reported experience of a research participant. For the latter, the path from a negative experience after drug exposure, use of a non-study medicine or a previous medical history, to their inclusion (or not) in a safety dataset, can be unexpectedly complex. Reports of subjective experiences (and indeed subsequent investigator assessments and interpretations) could be shaped by various factors, including the socio-cultural context of the study, individual perceptions of harm or what constitutes a medicine, or even specific personalities (56-58). Another potential influential element is the way participants are questioned to elicit this information. There have been advances in developing standard methods for measuring effectiveness outcomes (59). However, there has been less work relating to safety outcomes, aside from some tools developed for specific expected ADRs. Because there is no consensus regarding the best practices for how AEs in general should be elicited, individual study results can be difficult to interpret and compare, and the application of sophisticated statistical methods for meta-analysis is undermined (4, 60-62). This negates the likelihood of realising the necessary power to detect rare and latent effects in a real-world situation by pooling data from many sources, which would be of immense value for drugs used in understudied populations, such as the antimalarials. This thesis will, therefore, critically examine factors influencing the participant-reported data that are integral to the assessment of drug-related harm in resource-poor settings as this should inform appropriate questioning methods.

CHAPTER 2: LITERATURE REVIEW

2.1 Introduction

As discussed in the last chapter, a variety of interested parties monitor a drug's safety and tolerability profile throughout its development and beyond its registration. While serious ADRs are obviously critical to detect and evaluate accurately and timeously, less serious ADRs are also important contributions to associated morbidity and have the potential to negatively impact on adherence and treatment-seeking (48). Drug selection will also particularly hinge on an adverse effect profile where there are minimal differences between efficacies (63). This is important for 'relatively safe' drugs such as the antimalarials. Regardless of which study design is used, evaluating most ADRs is contingent to some extent on the accurate reporting by participants of their medical histories, subjective AEs and non-study medicines used. Indeed, often participants are the only source of these data. However, as will be shown in this review, there is a general lack of guidance about the methods for eliciting such information, and differing opinions on the optimal method(s). This reflects the historical marginalisation of measuring and reporting harms during drug development and post-registration (64-66).

Investigating the impact of elicitation methods on participant-reported safety data is complicated by some inherent methodological problems, largely related to the difficulties of measuring multiple possible subjective end points. Health and illness experiences are shaped by how people give meaning to them, and thus can be viewed as social, psychological and cultural constructions as well as biological realities (67-70). Medicines are also understood differently depending on underlying beliefs about issues such as their potential benefits or harms and the cultural context (71). Regardless of these influences, illness and the use of treatments need to be identified as accurately as possible in order for adequate clinical management. Often this relies on questioning patients in such a way that they can give a good description of their problem (and actions taken thus far) to healthcare workers. These interactions can be surprisingly challenging, leading one clinician to bemoan the inability of patients to express themselves clearly enough so that the symptom could be properly identified (72). Admittedly this was early in the 20th Century when it was only just being recognised that symptoms were not just indicators of disease but also expressions of a patient's experiences. This realisation, in fact, paved the way for the quality of life instruments we use now in both clinical practice and drug research (72).

As for healthcare communications in general, effective communication with study participants is vital for successful research, whether hospital, clinic or field-based. Over and above practical, technical issues, it is important to carefully train staff in communication skills relevant for the context and personnel involved (73). Formative qualitative work is important for clinical research, over and above its usefulness in sensitising communities to the research aims and practicalities and ensuring acceptability, it may also help in the development of culturally-appropriate questioning instruments. Some contexts may require specific tools to aid communication, such as the use of pictorial reporting methods in areas of low literacy, or pictures that convey information about health or the research in a sensitive and responsive manner to the local culture (74-76).

While there will be similarities between the reporting of health and use of treatments by patients in clinical practice and the reporting of experiences relating to drug effects in clinical research, there may also be differences. This issue for clinical research is not well studied. However, there is evidence that reports pertaining to ADRs are affected by issues such as personality, the cultural context of the experience and whether ill health is still viewed as such, a so-called response-shift (56, 58, 77-80). For clinical trials specifically, it has been proposed that fear of lack of confidentiality, stigma or negative repercussions may curtail AE reports, while nocebo effects (harmful response to an inactive product) may be related to a participant's expectations about the drug's effects and even information contained in consent documents (24, 81, 82). Meanwhile, conflicting interpretations of AEs between participants and researchers that manifest as discrepancies between the participants' experiences and eventual data fields will negatively impact study results (83). It is against this uncertain background that one may consider how participant-reported data relating to the assessment of harms in clinical trials and pharmacoepidemiology studies are gathered.

In order to contextualise and justify the planned research, I will first review the regulatory and other guidance documents regarding this topic, and then draw on available literature that considers the scope of methods used to elicit these key variables that are integral to the assessment of drug-related harm in clinical trials and pregnancy registries. I will explore the rationales behind the use of particular methods, including their influence on participant

reporting of these data. As will be seen, this is an under-researched area, especially so in resource-poor settings. Therefore, in the absence of literature specific to clinical trials or registries, relevant literature about participant-reported variables in pharmacoepidemiology studies in general will be reflected upon. I make the case that more work is needed to inform this important, but somewhat overlooked topic, both in general and relevant for malaria endemic and resource-constrained areas. This will set the scene for the 4 studies that make up this thesis. For the regulatory perspective, titles of all US Food and Drug Administration (FDA), European Medicines Agency (EMA) and South African Medicines Control Council (MCC) documentation relating to clinical drug development and pregnancy registries were considered. The full texts of those deemed relevant in terms of the elicitation of safety data were then examined in detail. The systematic review search methodology reported in Chapter 3 was also used to uncover other literature relating to the elicitation of relevant data in clinical trials. For literature pertinent to pregnancy registries, all websites of those registries listed by the FDA were examined, and the term “pregnancy registry” applied in PubMed. English full texts of those titles and/or abstracts that appeared relevant were then reviewed where available. The literature relating to health communications in general is broad, so to narrow the search for articles relevant for the clinical research context, as discussed above, a PubMed search using the terms "research" "communication" "Africa" was used. References of all selected literature were also examined and a search of articles by authors known to the candidate to be working in a relevant field was also made.

2.2 Requirements and guidelines relating to the elicitation of participant-reported safety variables

2.2.1 Clinical trials

The two largest regulatory authorities that regulate clinical trials, the US FDA and the EMA, as well as other agencies such as the South African Medicines Control Council (MCC), subscribe to the International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use (ICH)(84). The ICH general principles for the conduct, performance and control of clinical trials stipulate that "methods used to make the measurements of the endpoints, both subjective and objective, should be validated and meet appropriate standards for accuracy, precision, reproducibility, reliability, and responsiveness" (39). This suggests that all harm-related endpoints, including all participant-reported AEs, must adhere to this principle, which is not reflected in other

relevant documentation, is not being done in practice, and may indeed be unworkable. There is no mention in any FDA documentation of an absolute requirement to use a particular method for eliciting AEs. What is mandatory, however, is the need to be explicit about the elicitation method(s) in study reports: "the means of obtaining adverse event data should be described (volunteered, checklist, or, questioning), as should any specific rating scale(s) used" (55). This is so that reviewers evaluating safety data in licensing submissions may take the methods into consideration, because of their potential impact on the trial results (61).

When to solicit particular AEs

The FDA notes that "certain kinds of adverse effects are not likely to be detected or readily reported by patients without special attention" (85). An example is given of antidepressants that may have significant effects on sexual function, for which they suggest that "the use of targeted safety questionnaires or specific psychometric or other validated instruments is often important for such assessments, since routine adverse event monitoring and safety assessments tend to underestimate or even entirely miss such effects". Other clear recommendations for active solicitation of certain AEs are given in several of the therapy-specific guidance documents. These texts present the FDA's current thinking about various issues, though following them is not mandatory. One such guideline highlights the importance of directly asking participants about suicidal ideation and behaviour when investigating antidepressants or other drugs with the potential for this ADR, and recommends a particular scale (86). This is because a lack of complete case descriptions and an ill-defined baseline status of participants as regards suicide risk was deemed likely to be responsible for the lack of a prospective identification of such events previously. Again, drug developers are not beholden to this method and may choose another way "if the approach satisfies the requirements of the applicable statutes and regulations". A similar endorsement to actively solicit such AEs is made by the EMA (87). Likewise, for drugs with the potential for impairing driving, it is recommended that "subjective evaluation of CNS effects (e.g. by visual analogue scale) can contribute important information about the degree of subjective awareness of objectively demonstrated drug-related impairment. Both open-ended and targeted questions regarding adverse effects should be used" (88). A further example is the recommendation for specific solicitation of subjective reports relating to the potential for abuse if there is a perceived risk of harm for a particular drug (89). Ultimately it will be challenging for an

applicant to achieve registration in such circumstances should due attention to soliciting specific potential ADRs have not been paid.

Of note is that, when an instrument is used to elicit participant reports of AEs, it may not always be possible to present the data as frequencies (unless for instance a particular score signifies incidence of a specific AE). Instead, the data collected may represent a concept about health and, in these circumstances, are presented without interpretation from a clinician. The data may still be submitted to authorities to support claims made in medical product labelling provided they reliably measure the claimed concept in the population enrolled in the clinical trial (7). However, such patient-reported outcome (PRO) measures have traditionally been used for efficacy outcomes and their use for outcomes relating to harm is far less developed or understood (90).

Allowances for limiting the 'noise'

Another topic that impinges on the elicitation of data from participants that is addressed by regulatory agencies is the extent of AE and concomitant medication information necessary to collect at particular stages during drug development. In late phase trials, for a well-profiled drug being investigated in a known population, it may be acceptable to limit the data collected in order to reduce the burden on trial teams (91). For instance, investigators may consider only reporting ADRs or concomitant medications that will add information, provided expected ADRs, drug-drug interactions and metabolic pathways have been characterized. Where these data minimisation strategies overlap with concerns about the elicitation of participant reports is that trial staff may have differing opinions on which information reported by a participant needs to be recorded based on these criteria. Staedke et al. raises the possibility, for instance, that there may be room for a biased under-reporting of data if it reduces workload associated with completion of case record forms and other administrative requirements (92). This highlights the importance of carefully implementing a protocol in detail in a standardised manner, particularly in multicentre trials (11).

Other relevant guidance

Aside from the regulatory agencies, a Council for International Organizations of Medical Sciences (CIOMS) working group report emphasises the fact that the interaction and dialogue between clinical trial participants and investigators has not received much attention. Indeed they consider this to be "one of the most important issues that is rarely addressed" (4). The group gives clear recommendations that the process should be consistent site to site and program to program, and clearly outlined in protocols, informed consent documents and during training. They go a step further than the regulatory agencies and make a specific recommendation that "it is probably best to frame questions to the patients in general terms rather than invoke the possibility that study treatment may be responsible for ill effects". Care should therefore be taken to avoid leading questions. Furthermore, "although it is not advisable to read a specific list of possible ADRs when soliciting the patient's recent experience, patients should be alerted to known signs and symptoms indicative of medically important suspected or established ADRs in order to alert the investigator as early as possible". An example is given of muscle pain and/or tenderness in trials of HMG CoA enzyme reductase inhibitors (statins) that may be associated with rhabdomyolysis, i.e. a relatively non-serious AE that nonetheless may signal development of a more serious condition in time. They caution, however, that this type of "warning" should not be used routinely, and only under special circumstances. The CIOMS working group also goes beyond regulatory agency recommendations by advocating for specific enquiry of participants about their use of herbal and other non-traditional medicines. Furthermore, CIOMS specifically mentions the lack of regulatory guidance for the challenging task of gathering subjective data from young children, those with cognitive disorders such as Alzheimer disease and other scenarios where a proxy may need to be involved in a conversation about health.

A significant guide with relevance for this area of work is the extension of the CONsolidated Standards Of Reporting Trials (CONSORT) statement to include outcomes relating to harms (5). This initiative brings together trial methodologists, guideline developers, journal editors and research funders to make recommendations about how trial results should be reported. These have obvious bearing therefore on the way such trials are conducted. Specific to the elicitation of AEs, authors are asked to clarify in reports how harms-related information was ascertained in their trial, whether through the use of questionnaires or

interviews for instance, because of concerns about the impact on results of data being elicited, for example, through passive versus active surveillance.

Another scheme, the Core Outcome Measures in Effectiveness Trials (COMET) initiative, represents those interested in the development and application of core outcome sets (93, 94). There are less outputs relating to harm in this directory than for efficacy outcomes, again reflecting the general lack of attention to how harms data have been prioritised, obtained and reported. This will undoubtedly grow with time. Meanwhile, the methods for developing core measures for efficacy may be used as a guide for those wanting to do the same for harm (e.g. through use of the Delphi technique) (3, 95).

For malaria efficacy trials specifically, the WHO recommends that, to ascertain the incidence of adverse events (AEs), participants be asked about symptoms that have emerged since the previous follow-up visit by "direct questioning", though the detail of how this should be done is not provided (96).

2.2.2 Pregnancy registries

There is markedly less regulatory or other formal guidance about the conduct of pregnancy registries compared to clinical trials. However, the FDA does specify that they should be conducted in adherence with the Good Pharmacoepidemiology Practice Guidelines, which themselves specifically recommend the use of validated instruments and measures where they exist (97). In addition, there should be a detailed description in protocols of how data about health and exposures will be obtained, which may be directly from enrolled women via questionnaires or structured interviews. Plans for any pilot testing of data collection methods should be described and, where data are derived directly from participant reports, there should be validation of a sample of such reports, for instance against medical records (98, 99).

2.3 Participant-reported data elicitation methods used by researchers in practice

2.3.1 Clinical trials

As indicated above, there is an onus on researchers to describe their AE elicitation processes in study reports and publications. However, there is limited guidance about which methods to use, and when, and the guidance lacks the detail needed to enable consistency between studies. As will be seen, a wide range of methods are employed in practice, which can undoubtedly impact the data detected and compromise interpretations of the results and meta-analyses (100). It is beyond the capacity of this review to include all potential methods as they are countless. Publications also rarely present detailed information on the operational aspects of study conduct, particularly for drug safety data ascertainment, including for malaria trials (30, 101, 102). However, there has been notable work in some therapeutic areas to standardise aspects of collecting, assessing and/or reporting drug-related harms endpoints and a body of literature concerning the general debate about questioning participants to elicit AEs. No relevant literature concerning the elicitation of medical histories and non-study medication in clinical trials was found and limited work relevant to the elicitation of AEs in malaria clinical trials.

The usefulness of tools for measuring the often debilitating side-effects of drugs used in psychiatry has been recognised since the 1960s, explaining the abundance of instruments compared to other therapy areas (103). Such tools include standardised objective assessments by a trained healthcare worker (e.g. observation of tremor or gait), with some incorporating the elicitation of symptom reports from patients/participants. An early example is the Barnes Akathisia Rating Scale (BARS), which was developed to better assess the distressing motor restlessness side-effect induced by many antipsychotics that were previous overlooked or misdiagnosed (104). This validated tool helps patients put their symptoms into words through systematic probing, thereby allowing for discrimination of restlessness from other similar conditions, together with the clinical observations. Another early tool developed in recognition of the need to deeply probe for AEs was the Systematic Assessment For Treatment Emergent Events (SAFTEE), a semi-structured interview with over 70 items (105, 106). A more recent example is the self-administered Liverpool University Neuroleptic Side-effect Rating Scale (LUNSERS) (107). This assesses a comprehensive range of 41 potential expected neuroleptic ADRs relating to posture,

movement, vision, skin-related issues, mood and concentration. Included are 10 'red herring' items, unexpected symptoms that could highlight possible over-reporting, and thus help ensure validity. A limitation of this method, and others like it, is its reliance on literacy and adequate eyesight (103). Another scale is the Antipsychotic Non-Neurological Side-Effects Rating Scale (ANNSERS); a mixture of objectively-measured items and subjective reports that aim to, again, broaden the reach for potential expected ADRs, particularly with regard to sexual and urinary dysfunction (108). A patient-rated version of the well-known Udvalg for Kliniske Undersogelser (UKU) side-effect rating scale (UKU-SERS) is now also being used in clinical research, as is the Arizona sexual experience scale (ASEX), which is used to elicit reports of sexual dysfunction (109-111). For children specifically, Greenhill et al. found that more detailed question methods increased the number of AEs reported (112). Of note is that the clinicians found the detailed checklists took too much time, while parents found hem acceptable. Indeed, in recognition of the time needed to apply many of the above psychiatry instruments, simpler scales are now being developed for use in the clinic, though their application in research is not yet clear (113, 114). Despite the availability of so many instruments, a survey of 196 reports of studies conducted in this field found most were unclear about how ADR data had been obtained (63). Of those referencing a method, published rating scales were frequently used to measure movement disorder outcomes. However, there was far less consistency in, or clarity about, the methods used for eliciting non-neurological AEs, including sexual side-effects. This is surprising considering the established understanding that these are often underreported, and the recommendation by the FDA to explicitly assess sexual dysfunction (115).

The field of oncology has perhaps led the way as regards the standardisation of recording AE data. The pivotal National Cancer Institute's Common Terminology Criteria for Adverse Events (CTCAE), derived from the WHO Common Toxicity Criteria (CTC) developed in 1983, has been a mainstay of oncology clinical trials since 2003, and in fact is used in other therapy areas (116). An important distinction, however, is that this tool, which largely contains objectively evaluable signs or diagnoses, is itself not an AE elicitation method - there is no suggestion or evidence that clinical trial staff use it in a uniform way as a checklist against which to ask about symptoms (117). In fact, oncologists have been concerned for some time whether the data captured in AE reports accurately represent patient's experiences, reflecting a similar concern about post-marketing reporting of suspected ADRs (66, 83). For instance, there is evidence that clinicians may downgrade

patients' severity assessments and may even filter AE reports out entirely (118). To attempt to address this issue, a new online patient-reported outcome version of the CTCAE is being developed (119). It remains to be seen how this will be incorporated into clinical trials - whether the tool will be used to highlight patients' experiences prior to clarification by a clinician, or whether patients' AE reports will be listed in dossiers alongside the clinicians' reports.

Those working in the field of rheumatology clinical trials have also made a highly significant contribution to the development of standard approaches to measuring outcomes through the Outcome measures in rheumatology (OMERACT) collaboration (120). Emphasis has always been placed by this group on better representing the patients' perspective. To this end, a pivotal step forward was the finding that fatigue was an important outcome for patients that had not been recognised by researchers (121). While methods for coding data collected by investigators had been standardised, there was concern this did not extend to the elicitation, characterization, and severity grading of AEs (122). Since 1996 a group of interested researchers have thus been collaborating to address these issues. Initially, a rheumatology CTC instrument was developed for use by clinicians, though not an elicitation method per se. In fact, it was emphasized that it is not a checklist, and that investigators should use a standard open question approach to elicit participant reports for AEs) (122). However, as for oncology, a patient reporting tool is being developed, which could be used alongside the clinician's assessment (123).

The examples given above show that researchers, particularly in psychiatry, have long recognised that the method used to elicit AE reports directly impacts on the data collected, and hence the trial results. This has resulted in some standard tools being developed for AEs of special interest, but also for more general AE elicitation, in this field. In parallel, there has been debate in the literature about optimal method(s) for eliciting AEs in clinical trials in general, and pertaining to many other therapy areas (124-134). Despite the CIOMS recommendations that AEs should not usually be actively solicited, there are those who do advocate for a more systematic method, such as through the use of checklists of potential AEs. It has even been suggested that all trials use a generic checklist-type tool to elicit unanticipated effects (135).

The potential impact of the questioning method on the AE data collected has been investigated for at least 60 years. One of the first experiments was conducted in 1953 by Glaser and Whittow, not long after the first randomised controlled trial of streptomycin in 1946 (136, 137). In a class experiment they showed that questionnaires resulted in more symptom reports than when students were given blank forms to write on. The conclusion was that the questionnaires suggested, and even may have 'induced', symptoms. About 15 years later Avery et al. demonstrated in a patient clinical trial population that AE data will not be comparable between clinical research studies should researchers use different methods, although they were unclear about whether an open general question or questioning in response to a checklist of possible AEs produces more valid results (138). This was in light of their assumption that "suggestible patients will report symptoms when questioned, even though there is little objective evidence of the symptom being present". Over the years others have conducted similar experiments within, or relevant for, clinical trials, showing the increase in sensitivity of AE data detection by more detailed questioning (106, 112, 138-165). Reasons for recommending one questioning method over another, where given, fall into two broad categories: a) a preference for open, general, non-leading enquiries, as they prevent suggesting an AE to a participant, or b) a preference for more specific enquiry methods because they are more systematic, better able to be standardised, and will prompt for the elicitation of typically under-reported AEs, some of which may be unexpected. Over and above this, however, there are different opinions about the impact of the questioning method on the nature of the AE data elicited: which method, or combination of methods, elicits data that is more meaningful from a clinical or patient perspective, which is superior at detecting differences between study arms, or whether in fact different methods are complementary (141, 165, 166). A seminal editorial in 2006 by Ioannidis et al. called for more investigation about this topic (167). However, the studies to date have been diverse in terms of the methods compared, the populations and therapy areas in which they are embedded and the end points measured, including their definitions. This makes it challenging to understand the outputs as a whole. None of these studies were conducted in a resource-poor setting and none appear to have considered the impact of the elicitation method on the other variables necessary for assessing AEs in terms of whether they are likely to be ADRs. Recently a consortium of parties from industry, regulatory authorities, academia, the private sector and patient representatives interested in improving safety reporting in clinical trials (and other study designs) by better incorporating the perspective of the patient has been formed (168). The consortium has produced guidelines as to how patient-reported outcomes measures could be developed in

the same way that health-related quality of life instruments are, and some potential options as to how data could be presented in relation to AE data collected through traditional means have been proposed.

There has been limited attention to this topic in malaria research, however, an important recent contribution is a mixed method study that used a participatory approach to design novel forms for the collecting of AEs by non-clinicians in post-approval studies(74). Although the focus was on the healthcare worker's perspective of participants' reporting behaviours, this study highlights the importance of taking a patient-focussed approach to pharmacovigilance reporting. Thus, the field workers tasked with using forms were engaged in co-designing a pictorial storyboard to communicate the rationale for the information needed and facilitate rapport between the reporter and the respondent.

2.3.2 Eliciting participant-reported data in pregnancy registries

There has been no equivalent debate in the literature relating to the elicitation of AE or medication reports in pregnancy registries as there has been for clinical trial AE data. In addition, the detail of how participant-reported data are elicited in pregnancy registries is hard to find, although some limited information may be obtained from registry websites (169). The most common method appears to be structured interviews of registry participants by trained personnel, although pregnant women themselves may complete and submit a form, and may also be given a diary or note book in which they, or their health provider, may record relevant information (9, 43, 170, 171). While unclear, the rationale for these methods is most likely the methodological work reviewed by West el al, detailed below, including the frequently-cited work on optimal question design by Mitchell (172, 173). This showed that asking about drug use by specific indication, and by specifically named drugs, increased pregnant women's reports compared to open-ended questions; thus completeness of ascertainment is directly related to the specificity of the questions asked. As West et al. claim, the current best practice for soliciting exposure data from participants is to use memory aids, such as calendars and photos of drugs, although this is not apparent from publications (173). There is limited experience of prospective studies involving pregnant women in Africa, and it is unclear whether such methods are suitable for this context, and whether there are region/population-specific barriers to reporting data.

2.4 The validity of self-reported data

2.4.1 The pharmacoepidemiology literature viewpoint

West et al. note that answering a question relies on its comprehension and interpretation, search for and retrieval of information to construct an answer, judgement to discern the completeness and relevance of memory for formulating a response and development of the response based on retrieved memories (173). These attributes are thus important to consider when formulating questions. Validating participant reports is particularly hampered by the lack of an adequate gold standard as there is often no other solid record of a symptom, diagnosis or medicine use (128). However, studies have shown that for medicines used in pregnancy (derived from both prospective cohorts and studies that enrol women after a birth defect is diagnosed), recall is negatively affected the longer the time since drug exposure. In addition, the acute use of a medicine was less easily recalled compared to one used chronically (173). For drug-use questioning in general, medication- and indication-specific questions are more sensitive compared to general questions, while even more are reported through the use of calendars or photos of drugs during interviews (173, 174) .

There appears to be better recall of specific, familiar illnesses such as asthma and diabetes, compared to less well understood, but equally common conditions, like sinusitis and back pain, and accuracy may be affected by participants' willingness to report (173). For instance, sexually transmitted illnesses and mental disorders may be underreported due to stigma, embarrassment about reporting or fears about loss of confidentiality. Illnesses that were reported more accurately included those that were more recent, had a greater impact on a person's life, and required repeated medical visits or hospitalisation (more so if surgery was involved). The evidence for whether demographic or other factors influence the data elicited is conflicting and/or affected by study design issues, though there may be a positive effect if patients are better motivated to answer questions. In addition, as for exposures, the questioning method is influential; checklists improve reporting, particularly if simple questions are used as these help respondents understand what information is required (173).

There is a paucity of research regarding the validity of participant reported safety data in resource-poor regions. However, Hodel et al. found significant underreporting of prior use

of antimalarials by patients being screened for enrolment in clinical trials (25). There is also evidence of underreporting of prior and concomitant non-study medication by participants who were co-enrolled in two microbicide clinical trials (175). Specific to malaria, even clinical research teams who have standardised their safety data collection strategies report differences in AE frequency or nature between studies and sites; they postulate this may be related to trial methodology (92). It is likely that there will also be methodological differences between how other research teams capture these endpoints.

2.4.2 An alternative view of validity

The pharmacoepidemiological literature on the validity of self-reported data summarised above suggests a positivist view (epistemology); that there is a true health issue which should be measurable through scientific methods (176). The gold standard against which this truth is compared varies depending on the researcher's opinion as to which measure is the most likely to be valid (e.g. a hospital chart, prescription record, responses to a questionnaire or otherwise). An alternative to this is view is to take a social-constructivist epistemology; the health experience of an individual is constructed by humans and may depend on social, historical or other contexts (177). Participants' responses may therefore be reflections of their realities within a particular time, place or context, and the experience may be fluid and subject to interpretation by the various actors involved. This corresponds with views on the body moving away from it being a fixed physiological entity to considering that society and other factors could be influential (178). Observations of, and ideas about our bodies as we move through a myriad fluctuating physical, cultural, social and political environments have resulted in various notions of what has been termed embodiment (179). Attempts to investigate health (however one may define it) is likely therefore to be enriched by considering a more holistic approach compared to investigations that focus on a more limited objective assessment. This is important to consider when designing strategies for health-related communication.

2.5 Conclusion

A lack of evidence about the factors influencing the elicitation of participant-reported data, including guidance on methods for questioning participants in clinical drug trials and pregnancy registries, may lead to less than optimal reporting of medical histories, AEs and non-study medications. This will negatively affect individual studies, as these reports

underpin subsequent steps for representing an experience as an eventual safety end point. Meanwhile, disparate methods between studies will impede data synthesis and comparisons. While this issue is pertinent to all drug studies, there is a particular lack of relevant methodological work relating to resource poor settings, and specifically the effects of antimalarial drugs. This is of greater importance now, when efforts to improve malaria control and eventually eliminate malaria have led to these treatments not only being given to symptomatic malaria patients, but also as preventive treatment to those 'at risk' but who do not have malaria (29). The areas where antimalarial drugs are being distributed widely have generally less well advanced regulatory systems so vulnerable groups, including pregnant women and those with co-morbidities such as HIV, may be exposed. It is important therefore to investigate elicitation methods for participant reporting of safety related data in these contexts. Better knowledge about these issues should assist antimalarial and other drug researchers to design and test novel strategies to overcome barriers to achieving accurate safety data in similar contexts earlier in drug development. Understanding more about the factors that influence participant reports of AEs and other relevant data in antimalarial drug research will support efforts to perform pooled analyses and also clarify the safety profile of drugs introduced into healthy but at-risk populations within malaria control and elimination strategies.

2.6 Gaps in the evidence that will be addressed by this thesis

Despite the relative lack of relevant methodological work in antimalarial clinical trials and pregnancy registries, it is useful to reflect on the work conducted in other therapeutic areas, as summarised above, when considering the way forward for these fields. However, equally, it will also be important to consider whether there are any likely specific, contextual, influences that set antimalarial clinical trials and pregnancy registries apart from the currently available literature. Due to the varied studies that have compared AE elicitation methods relevant for clinical trials there is a need to explore this research in a systematic way. A Cochrane systematic review will therefore be conducted about this topic in the thesis. The review itself, meanwhile, would benefit from inclusion of a study from a resource-poor environment due to a lack of relevant work in that context. Such a study was therefore designed to nest within antimalarial-antiretroviral interactions clinical trials. As none of this type of work to date has explored with the participants themselves why they answer questions differently, qualitative methods to investigate such reporting behaviours were also included. Given the high burden of malaria carried by pregnant women in Africa,

qualitative work was also nested within a pilot pregnancy registry, in order to understand any specific challenges in eliciting accurate exposure reports from pregnant women enrolled in such a registry in Africa. Finally, as elicitation methods are rarely described in publications, it would be of value to survey antimalarial drug researchers about the methods they use to elicit these data and why. Together these investigations should contribute to the wider debates about optimal methods for eliciting safety data in clinical trials in general and also provide some much needed evidence relating to these methodological considerations specific for resource-poor regions. This work will pave the way for finding appropriate questioning methods for use in these contexts.

2.7 Aim and objectives

The overall aim of this thesis is to critically examine factors influencing the participant-reported data that are integral to the assessment of drug-related harm in resource-poor settings. This will enable recommendations for the future as regards the design, conduct and/or reporting of studies collecting participant-reported safety end point data, and further methodological research required. Such recommendations should strengthen AE reporting methods. Specific objectives, that will be justified in section 2.9, include:

- To systematically review available literature on research that compares the methods used within or relevant for, clinical drug trials/studies to elicit information from the participants about AEs (**Chapter 3**).
- To survey the methods used to detect, assess and record participant-reported medical history, AE, study drug adherence and non-study drug data in anti-malarial drug clinical research (**Chapter 4**).
- To investigate the factors shaping reports of medical histories, AEs and previous or concomitant non-study medication data by participants who had malaria and/or HIV and were enrolled in antimalarial and antiretroviral drug interaction clinical trials (**Chapter 5**).
- To explore, from pregnant women's perspectives, barriers to their reporting of exposure data in a pilot pregnancy registry and how these might be overcome (**Chapter 6**).

2.8 Choice of methods and research context

2.8.1 Methods and research position

The research will use primary quantitative and qualitative study designs, survey and secondary data synthesis techniques, as all these methods have the potential to contribute to understanding the rationale for questioning methods and reasons for particular participant safety data reports in clinical trials and pregnancy registries. As mentioned in Chapter 1 the focus on these particular designs were for pragmatic reasons. The narrative systematic review reported in Chapter 3 and the survey reported in Chapter 4 set the scene for two pivotal, primarily qualitative, studies for this thesis that are intended to better comprehend behaviour and social interaction in drug safety research. While people may not always believe they can explain what they experience, why they behave in a certain way or how they consider their role in drug research, rigorous social science methods may assist in providing evidence of socio-cultural factors affecting what is primarily considered a quantitative field (180). However, there is a tension between the candidate's background and role in pharmacoepidemiology (which is largely positivist in outlook as indicated in Chapter 2) and the qualitative methods required in order to contribute to the achieving the aims of this thesis necessitating a social-constructionist perspective. This tension is explored within the discussion. Further justification for the specific choice of research methods used in each chapter is given in section 2.11 below.

2.8.2 Research context

A significant part of this work will be nested within studies conducted by the UCT-MRC Collaborating Centre for the Optimisation of Antimalarial Therapy (CCOAT) in understudied populations in Africa. Three studies were supported by the ACT Consortium, which is funded through a grant from the Bill and Melinda Gates Foundation to the London School of Hygiene and Tropical Medicine (37). The ACT Consortium is a global research partnership of a number of eminent public health and academic institutions in Africa, Asia, Europe and the United States working on over 25 projects in 10 countries to answer key questions on malaria drug delivery. The Consortium uses a variety of methods to address research questions to optimise the use of artemesinin-based combination therapies (ACTs). Access to ACTs is difficult for many people who need them, while others who do not need them (e.g. they do not have malaria) are exposed unnecessarily. The ACT Consortium is working to investigate why this is so, what the implications are and how the situation may be

improved. Including safety as a theme was important for the Consortium as, despite the wide scale use of ACTs, little is known about their long-term effects and their use in some vulnerable populations, such as those co-infected with HIV. Moreover, it is clear that there is a need to work towards the pooling of safety data from multiple studies to understand uncommon adverse drug reactions that may not be adequately characterised by individual studies. As such, all ACT studies had the potential to contribute data to a safety data repository. In order to understand more about factors influencing the safety assessments of ACTs, the candidate, who is a senior clinical research manager at CCOAT, nested work included in this thesis within two ACT Consortium artemether-lumefantrine and antiretroviral drug interaction studies, conducted a global survey of medical history, AE and non-study medication data collection methods and a systematic review of participant questioning methods used to collect this data.

While none of the ACT Consortium studies enrolled pregnant women, there was a valuable opportunity to nest relevant work within a pilot WHO Pregnancy Registry study (43). The registry was designed to build capacity for obtaining reliable information on obstetric, medical and drug histories during pregnancy in resource-poor settings. This will help quantify both the baseline risk for major congenital malformations and the risk associated with various exposures. These are particularly important for the artemisinins which, although generally considered safe in humans, animal studies suggest they may have the potential to cause foetal loss and congenital malformations if taken during early embryogenesis (181). As pregnant women are typically excluded from clinical trials by a manufacturer there is a lack of evidence for antimalarial safety in this vulnerable population. Meanwhile, women of childbearing potential or already pregnant have access to myriad medicines, including antimalarial drugs, over the counter. Exploring the factors that influence pregnant women's reports of exposures was important as these may differ from what non-pregnant clinical trial participants experience.

Ethical approval for the ACT Consortium primary research study and survey, and the WHO Pregnancy Registry analysis was obtained from the UCT Health Sciences Faculty Human Research Ethics Committee (HREC Refs: 376/2009, 103/2012 and 760/2013). Approval was also obtained from the National Institute of Medical Research, Tanzania for the Tanzanian part of the ACT Consortium primary research study. For the WHO Pregnancy Registry,

ethical approval was also obtained from the committees of the respective institutions or health authorities (Marekere University Faculty of Medicine, Moi University Institutional Research and Ethics Committee and the Ghana Health Service Ethics Committee). These approvals are presented in Appendix 1 together with the informed consent documents for each study in Appendix 2.

2.9 Thesis outline

Chapter 1 was an introduction to the fields of pharmacovigilance and pharmacoepidemiology, and their important role in resource-poor settings. It summarises some inherent methodological limitations in detecting both serious and less serious ADRs and the relevance of participant reported information. The potential for influence on these data from various factors, including the way participants are questioned about their health and use of medicines, was highlighted. In addition it suggested that it is important to optimise the accuracy of other key variables that contribute to the assessment of ADRs; medical histories and non-study medications.

This chapter was a review of the literature pertaining to the methods used to elicit participant-reported AEs, medical history and non-study medications in clinical trials and pregnancy registries, and the factors that influence the validity of reports in pharmacoepidemiology studies. Because the latter is merely one worldview on validity or the 'truth' about a health-related experience, a different perspective was given about ideas of health in relation to bodies within fluctuating physical, cultural, social and political environments. The literature review revealed some gaps in the research that would be useful to explore in this work. In particular, it would be useful to 1) conduct a systematic exploration of the studies that have compared AE elicitation methods, 2) clarify which methods are used to elicit such data reports in malaria clinical trials, and 3) the rationale for the choice of such methods from the researchers' perspectives. There is a distinct lack of relevant work in resource-poor settings, including in pregnant women who are often excluded from clinical trials. Incorporating qualitative methods into clinical trials and pregnancy registries will be invaluable to understand reporting behaviours from the participants' perspectives.

Chapter 3 is a Cochrane systematic review of studies that compared methods to elicit AE reports from participants within or relevant for clinical trials. A search strategy was designed to uncover studies that compared 2 or more AE elicitation methods: relevant data were extracted and then subjected to a narrative synthesis. This review was important to include in order to establish some evidence for how different questioning methods impact on the number and nature of the AE data elicited. A better understanding of the methodological challenges should inform potential solutions for achieving better, or more harmonised, AE ascertainment in clinical trials. The review was not, however, limited to antimalarials or resource-poor settings as it was clear that there were too few such publications. Nevertheless, it was also important to consider the wider pharmacovigilance environment when investigating this topic, as much of these method-effects are not disease specific and will be applicable for both malaria and other diseases.

Chapter 4 is a global online survey with clinical researchers about the methods used to elicit, assess and record AEs, medical histories, and non-study medication reports from participants taking part in antimalarial drug studies. Questions were designed in such a way as to elicit as much detail as possible about the rationale and application of methods to obtain such data in various study designs and populations. It also explored what researchers consider to be optimal (relevant, important and feasible) methods to elicit these data. Closed question responses are presented quantitatively and open questions as themes. This survey was intended to give a broad overview of the scope of methods being used in studies that may contribute AE data to pooled analyses of antimalarial drugs. In addition, the survey would help with understanding researchers' perspectives on the choice of elicitation methods as this would inform recommendations for improving conversations with participants to collect safety data. Exploring the processes for evaluating harm in anti-malarial drug research in this way may offer insight into how study participants' experiences become facts in databases.

Chapter 5 is mixed-method study nested in 2 African antimalarial-antiretroviral drug interaction trials. These trials both involved artemether-lumefantine but were conducted in quite distinct populations - otherwise healthy HIV positive participants in Cape Town, South Africa, and patients with or without HIV presenting with malaria symptoms at a district hospital in Tanzania. This provided a good opportunity to explore similarities and

differences between these contexts relating to the elicitation of participant-reported safety assessment variables. Two types of questioning were used consecutively to elicit medical histories, prior and concomitant non-study medications and AEs. Any participant who reported differently to the two questions was invited to an in-depth interview to investigate any further relevant data reports, and the reasons for their reporting behaviours. Interviewed participants were then also invited to a focus group discussion, where the reasons for differential reporting were further explored as a group.

Chapter 6 is a qualitative study nested within a pilot pregnancy registry study conducted in three countries in East and West Africa. Women enrolled in the registry and those from its source communities were invited to focus groups discussions to discuss their health seeking behaviour, which medicines and other substances they take while pregnant, and how they experienced reporting such exposures at the antenatal (ANC) clinic.

Chapter 7 provides a summary and discussion of the findings of the thesis and considers the implications of the findings on the evaluation of drug safety. It also recommends a way forward for optimising participant-elicited safety reports in future antimalarial drug trials and pregnancy registries.

CHAPTER 3: COCHRANE SYSTEMATIC REVIEW:

ELICITING ADVERSE EFFECTS DATA FROM PARTICIPANTS IN CLINICAL TRIALS

Elizabeth Allen[1], Clare Chandler[2], Nyaradzo Mandimika[1], Cordelia Reddy[1], Karen I Barnes[1]

[1] Division of Clinical Pharmacology, Department of Medicine, University of Cape Town, Cape Town, South Africa.
[2] Department of Global Health & Development, London School of Hygiene & Tropical Medicine, London, United Kingdom

Publication status:

The protocol for this review was peer reviewed and published by the Cochrane Collaboration in 2013:

Allen EN, Chandler CIR, Mandimika N, Barnes KI. Eliciting adverse effects data from participants in clinical trials (protocol). The Cochrane Library 2013;9.2.

The review presented in this chapter is ready for submission to the Cochrane Collaboration.

Author contributions

The candidate (EA) conceived of the review, wrote the protocol, conducted the analysis, and wrote the paper. CC had input into the protocol development. EA, NM and CR independently screened titles, abstracts or full texts for eligibility and KB arbitrated over any decisions about selection. CR performed a 100% check of data extracted. KB conducted an independent risk of bias review and supervised the writing of the paper.

Note about citations/referencing:

The first time each included paper is mentioned in the results section it is referenced in the thesis bibliography. Thereafter, the Cochrane style is maintained, whereby the referenced paper is indicated as (primary author date) within the text.

Abstract

Background: The analysis of drug safety in clinical trials involves assessing adverse events (AEs) either on an individual case basis or by aggregate statistical synthesis to provide evidence of likely adverse drug reactions (ADR). While some AEs may be ascertained from physical examinations or tests, there is a great reliance on reports from the participants to detect subjective symptoms, where the participant is often the only source of information. There is no consensus on how these reports should be elicited from participants, although it is well known that methods involving direct questioning influence the extent and nature of the data detected. This leaves room for measurement error, and undermines comparisons between studies and pooled analyses. This review investigated comparisons of methods used in clinical trials to elicit participant-reported AEs, such as a general enquiry, checklist, diary, memory aid etc., whether applied face-to-face or otherwise. This should contribute to knowledge about the methodological challenges, and possible solutions, for achieving better, or more consistent, AE ascertainment in clinical trials.

Objectives: To systematically review available literature on research that compares the methods used within, or specific for, clinical drug trials/studies to elicit information from the participants about the AEs that were defined in the protocol or in the preparation for the trial.

Search methods: Databases searched included: EMBASE (OVID), MEDLINE (OVID), MEDLINE in Process and Other Non-Indexed Citations, the Cochrane Methodology Register (Wiley Online), the Cochrane Central Register of Controlled Trials (Wiley Online), the Cochrane Database of Systematic Reviews (Wiley Online), the Database of Abstracts of Reviews of Effects (Wiley Online), the Health Technology Assessment database (Wiley Online), CINAHL (EBSCO), CAB Abstracts (OVID), BIOSIS (Web of Knowledge), the Science Citation Index (Web of Knowledge), the Social Science Citation Index (Web of Knowledge), and the Conference Proceedings Citation Index – Science (Web of Knowledge). The search was designed using thesaurus headings and synonyms for each of the following concepts: [A]: Adverse events AND measurement; [B]: Participants AND elicitation (also other synonyms for the extraction of information about adverse effects from people); [C]: Participants AND checklists (also other synonyms for the methods used to extract information about adverse effects from people). Pragmatic ways were used to limit the search results whilst trying to maintain sensitivity. There were no date, sample size or language restrictions in the searches. However, only English reports were included fully in the review, because of resource constraints as regards translation.

Selection criteria: Two types of studies were included - clinical drug trials that included a comparison of two or more methods within- or between-participants to elicit participant-reported AEs, and research studies performed outside the context of a clinical drug trial to compare such methods which could be used in clinical trials (evidenced by reference to such applicability). Primary outcome data included AEs elicited from participants taking part in any such clinical trial, however as terminology and definitions for AEs were unclear or inconsistent in many studies we included any participant-reported data relevant for an assessment of drug-related tolerability or harm, using the authors' terminology (and definition where available), with comment on whether the data were likely to be treatment-emergent AEs or not.

Data collection and analysis: Titles and abstracts were independently reviewed for eligibility. Full texts of potentially eligible citations were then sought and independently reviewed for final eligibility. Relevant data were extracted onto a pre-tested form and subjected to a 100% check. Disagreements were resolved by discussion, involving a third author as necessary. The risk of bias was independently assessed by two review authors. The Cochrane Collaboration's 'Risk of bias' tool was used for reports that compared outcomes between-participants, while reports that compared outcomes within-participants, each study was critically evaluated in terms of the potential impact of the design and conduct on the findings within the framework of potential selection, performance, detection, attrition, reporting and other biases. An attempt was made to contact study authors whose email addresses were available to retrieve protocols or specific relevant missing information. Reports were not excluded from this review on the basis of quality unless data for the outcomes were not possible to compare. As a quantitative meta-analysis of included studies was not possible due to differences in study design and presentation, a narrative synthesis was conducted.

Main results: Thirty-three studies eligible for inclusion in this review largely compared open questions with checklist-type questions or rating scales. Two included participant interviews. Despite different study designs, populations and details of questioning methods, the review showed that more specific questioning of study participants leads to more AEs being detected compared to a more general enquiry. A subset of six studies suggests that more severe, bothersome or otherwise clinically relevant AEs were reported when an initial open enquiry was used while some less severe, bothersome or clinically relevant AEs were only reported with a subsequent specific enquiry. However one study showed that quite severe or debilitating AEs were only detected by an interview, while others did not find a difference in the nature of AEs between elicitation methods. No

conclusions could be made regarding the impact of question method on the ability to detect a statistically significant difference between study arms.

Authors' conclusions: This review supports concerns that the methods used to elicit participant-reported AEs influence the detection of these data. There is the risk for under-detection of AEs in studies using a more general elicitation method compared to those using a more comprehensive elicitation method, some of which may be important from a clinical perspective or for patients themselves. This could compromise the ability to pool such data. However the impact on the nature of the AE detected by different methods is unclear. The wide variety and inherent low quality of methods used to compare elicitation strategies in clinical drug trials limited this review. Future studies would be improved by using and reporting clear definitions and terminology for AEs (and other important variables), the frequency and time period over which they were ascertained, how they were graded, assessed for a relationship to the study drug, coded and tabulated/reported. While the many potential AE end points in a given trial may preclude the development of general AE patient-reported outcome measurement instruments, much could also be learnt from how these employ both quantitative and qualitative methods to better understand data elicited. Any chosen questioning method needs to be feasible to use for both staff and participants.

3.1 Plain language summary

Questioning clinical trial participants about their health in order to collect adverse drug effects

Clinical drug trials or studies are usually conducted to assess how well the drug works but also whether it causes any harm (side effects or adverse effects). Adverse effects can be detected by the trial doctor examining participants or taking some blood samples or other kinds of tests. The trial staff can also ask participants about how they are feeling since taking the trial drug. However, the way participants are asked about their health can vary from trial to trial, or even within a trial. In one trial participants may be asked a simple open question such as 'how have you been feeling?', while in another participants may be asked about whether they have had any of a long list of possible symptoms (like 'have you had a headache, stomach ache, sore muscles etc.?). There has been concern that these different kinds of questions and how they are phrased will impact on what participants report about their health during a trial. This could then potentially affect the results and what we know about the side-effects of drugs. This review was therefore done to look at

studies that compared different types of participant questioning methods. The authors found 33 studies comparing mainly open questions with checklist-type questions, but also some ratings scales and participant interviews. While the studies were all very different in terms of the types of disease, drug and patients studied, it was found in general that, as would be expected, when a more specific type of question was asked (like a checklist) participants reported more symptoms. What is interesting is that, in those studies that looked more closely at the types of symptoms reported, it seems to be that an open question picks up the more severe or bothersome symptoms compared to a checklist-type question. However, some studies found that even quite severe or bothersome symptoms are not reported when a participant is asked an open question and these severe symptoms will only be reported with the more specific question. Therefore it is difficult to say whether one method is better than the other and the different questioning methods may, in fact, be complementary and therefore should be used together. It is, however, also difficult to say what a specific question should include as it can take too long to answer a very long list. While more work needed, it is very important for trials to be clear about which kind of questioning was used when they publish their results. This helps readers understand the findings about the side-effects and make comparisons between trials more accurately.

3.2 Background

3.2.1 Description of the problem or issue

Manufacturers must demonstrate safety, efficacy and quality of their investigational drug by way of clinical trials in order to achieve registration with regulatory authorities. Thereafter, they, and other stakeholders, continue to evaluate the product's harm profile in subsequent trials, particularly in under-studied population groups (44). Safety analyses in clinical trials largely involve identifying untoward medical occurrences after exposure. These endpoints, which are not necessarily causally related, are called adverse events (AEs) (1). AEs are assessed either on an individual case basis or by aggregate statistical synthesis to provide evidence of likely adverse effects (adverse drug reactions, ADRs), those AEs that have a reasonable possibility of being caused by the trial drug (4).

The processes involved in collecting, recording, analysing and reporting AEs are generally considered more complex than those involved in evaluating efficacy, and methods are relatively less developed (182). While some AEs may be ascertained from physical

examinations or tests, there is a great reliance on reports from the participants to detect subjective symptoms, where the participant is the only source of information. There is no consensus on how these reports should be elicited from participants, although it is well known that methods involving direct questioning influence the extent and nature of the data detected (61). For instance, studies have found that giving participants a checklist of potential AEs yields more reports compared with posing a general enquiry about change in health (139). However, it is uncertain whether one way of questioning over another is better for detecting ADRs (165). Should methods to elicit AEs be less than optimal, there is a margin for measurement error which will undermine individual trial results and meta-analyses of multiple trials. The latter will also be so should trials use disparate methods, restricting the ability to detect rare ADRs and explore factors influencing the assessment of harm (61, 182). This situation is compounded by generally poor reporting in subsequent publications about which methods to determine participant-reported AE data were used (5).

3.2.2 Description of the methods being investigated

This review investigated any method used in a clinical trial to elicit participant-reported AEs, such as a general enquiry, checklist, diary, memory aid etc., whether applied face-to-face or otherwise. Due to the lack of consensus as described above, the details of all methods studied were only known once the review was ongoing. Eligible studies included a comparison of methods used to elicit information on other participant-reported variables (e.g. concomitant medications or medical histories), were also included in the review.

3.2.3 How these methods might work

Little is known about how different methods of elicitation work. A study that aimed to identify barriers to accurate and complete reporting of harms data using qualitative methods suggests that questioning detail and terminology influences participants' recognition of health issues and treatments. Moreover, the authors have suggested that the perceived relative importance of health issues and treatments to the participant may be a factor (183).

3.2.4 Why it is important to do this review

Current heterogeneity in, and uncertainty about, the best practices for participant-reported AE elicitation in clinical trials leaves regulatory authorities, policy makers, healthcare professionals, patients and the public unsure how far results are accurate and comparable.

It would therefore be useful to synthesise research that compares elicitation methods. This should contribute to knowledge about the methodological challenges, and possible solutions, for achieving better, or harmonised, AE ascertainment in clinical trials.

3.3 Objectives

To systematically review available literature on research that compares the methods used within, or specific for, clinical drug trials/studies to elicit information from the participants about the AEs that were defined in the protocol or in the preparation for the trial.

3.4 Methods

3.4.1 Criteria for considering studies for this review

Types of studies

- Clinical drug trials that include a comparison of two or more methods to elicit participant-reported AEs
- Research studies that have been performed outside the context of a clinical drug trial to compare two or more methods to elicit participant-reported AEs, and which could be used in clinical trials (evidenced by reference to such applicability)

Types of data

AEs elicited from participants taking part in any such clinical trial. For the purposes of this review, AEs are defined as those outcomes that were pre-specified as potential AEs to be investigated in the trial (including expected or unexpected AEs, the latter which will not be known, but are intended to be detected during the trial/study), recognizing that the trial itself might reveal that these are not actually increased in the intervention group compared with the control group. Concomitant medication and medical history data were also included in this review, where eligible studies also included a comparison of methods used to elicit those. This is because these variables also impact on the assessment of whether an AE is likely to be an ADR, as well as their impact on the eligibility of participants for the trial. It became apparent during the review that terminology and definitions used for AEs were unclear or inconsistent. This is partly due to changing perspectives on this topic over time and partly because we included research studies outside the context of a clinical drug trial in the review. Thus we included studies that reported participant-reported data relevant for an assessment of drug-related tolerability or harm, using the authors' terminology (and

definition where available) with comment on whether the AEs were likely to be treatment-emergent or not.

Types of methods

Any combination of elicitation methods compared within- or between-participants. This included, but was not limited to, unstructured or structured enquiries, checklists or questionnaires (e.g. by body system, symptom etc.), diaries and memory aids.

Types of outcome measures

Primary outcomes

- The effect measure (or number, proportion) and/or nature (e.g. characteristics, severity, causality assessment) of AEs identified by the method of elicitation, as defined by the original authors.

Secondary outcomes

- If relevant, the effect measure (or number, proportion) and/or nature (e.g. characteristics, severity, causality assessment) of AEs identified by the method of elicitation by the relevant trial/study interventions
- If relevant, the effect measure (or number, proportion) and/or nature (e.g. medication class) of concomitant medications and/or medical histories identified by the method of elicitation, as defined by the original authors
- If relevant, summary results of qualitative methods used
- If relevant, results of inherent elicitation method validation studies.

3.4.2 Search methods for identification of studies

There was no date or sample size restriction in the searches. However, only English reports were searched for an included in the review, because of resource constraints as regards translation.

Electronic searches

The searches were designed and conducted with the assistance of an experienced information professional. A list of databases and search strategies was finalised prior to starting the search, with subsequent iterations fully documented. The following databases were searched: EMBASE (OVID) 1980 to 2015 week 11; MEDLINE (OVID) 1946 to March week 2 2015; MEDLINE in Process and Other Non-Indexed Citations, March 16th 2015;

Cochrane Methodology Register (Wiley Online) Issue 3 of 4, July 2012 (no longer updated); Cochrane Central Register of Controlled Trials (Wiley Online) Issue 2 of 12, February 2015; Cochrane Database of Systematic Reviews (Wiley Online) Issue 3 of 12, March 2015; Database of Abstracts of Reviews of Effects (Wiley Online) Issue 1 of 4, January 2015; Health Technology Assessment database (Wiley Online) Issue 1 of 4, January 2015; CINAHL (EBSCO) 1981 to March 2015; CAB Abstracts (OVID) 1973 to 2015 Week 10; BIOSIS (Web of Knowledge) 1969 to July 2013 (can no longer access); Science Citation Index (Web of Knowledge) 1970 to March 2015; Social Science Citation Index(Web of Knowledge) 1970 to March 2015; Conference Proceedings Citation Index – Science (Web of Knowledge) 1990 to March 2015.

The search was designed using thesaurus headings and synonyms for each of the following concepts: [A]: Adverse events AND measurement; [B]: Participants AND elicitation (also other synonyms for the extraction of information about adverse effects from people); [C]: Participants AND checklists (also other synonyms for the methods used to extract information about adverse effects from people)

Ideally the search would have been run using the following search string: A AND (B OR C). Unfortunately this produced unmanageable volumes of results, largely because it was impossible for the search to differentiate between (i) studies aiming to compare two different methods for eliciting adverse effects data (i.e. the included studies in this review); and (ii) studies which mentioned in the abstract that they collected data from participants about adverse effects (i.e. thousands of studies irrelevant to this review).

The information specialist, with help from information colleagues, used some pragmatic ways to limit the search results whilst trying to maintain sensitivity in the search. Each of these were used separately and then combined with OR: Frequency searching: retrieving only those records which contained certain adverse effect-related terms at least three times in the abstract, the rationale being that if the study is based on the collection of adverse effects data then associated terms would be used at least three times in the abstract. This part of the strategy was tested with different proximities (i.e. two times or four times) by comparing a sample of results from each to see what was being lost as the number increased; Title field: one part of the search retrieved only those records with adverse effects terms in the title, the rationale being that if the study is focused on collection of adverse effects data then associated terms would be in the title. This could

only be used in the databases accessed through OVID because other databases do not provide this functionality. Some of the other databases provided manageable numbers of results without these techniques; with others the title field technique was used.
There were no date limits on the electronic searches, but the searches were limited to the English language.

See Supplementary File 3.1 for the exact search strategies used in each database (presented with the results). Endnote X3® was used to collect, de-duplicate, screen titles/abstracts, and record decisions on inclusion of papers.

Searching other resources

The electronic searches were supplemented by a review author checking reference lists of included reports, some excluded reports that were relevant to the topic, and relevant reports known to the authors who are familiar with the research area (184), hand searching of recent relevant topic-area conference abstracts (International Conference on Pharmacoepidemiology and Therapeutic Risk Management, International Society of Pharmacovigilance annual meeting) (185), and searching online libraries of theses/dissertations.

3.4.3 Data collection and analysis

Selection of studies

EA examined titles and, where available, abstracts of identified citations in order to remove obviously irrelevant reports (e.g. non-human studies). Thereafter, two authors (EA and NM or CR) independently reviewed the remaining titles and abstracts for eligibility according to the Criteria for considering studies for this review in this protocol. The full texts for all reports that appeared relevant were sought, as well as those for which the title and abstract was insufficient to determine eligibility. Reports from the same piece of research were linked together. The same review authors independently assessed final eligibility, with disagreements resolved by discussion (involving a third author, KB, as necessary). While the review authors were blinded to each other's assessments, they were not blinded to any information in the titles, abstracts or full texts. All documents relating to this search and selection process, including the primary reason for non-inclusion, were recorded.

Data extraction and management

EA extracted data onto a data extraction form according to a pre-specified list, with a second review author (CR) checking 100% of fields. Disagreements were resolved by consensus, with, if necessary, a third author, KB, consulted to resolve disagreements. The original planned list was pre-tested in two reports and modified before being finalised as:

- Authors
- Date published
- Summary description of study methods including any drug(s), indications/inclusion criteria, assessments(s), and duration of follow-up
- Data (AE or equivalent with primary authors' terminology and definition, where available; medications and medical histories if these were also outcomes of the comparison)
- Comparisons (within- or between-participants)
- Elicitation methods, including (if available) description of their development and application methods. Also training/experience of staff, how AEs were described, whether verbatim reports were captured, and language.
- Outcomes and results
  - The relative effect estimates derived from one method of ascertainment versus the other, by study arm if relevant
  - The number/proportion and/or nature of AEs as defined by the authors of the original study, by study arm if relevant
  - If relevant, the relative effect estimates/number/proportion and/or nature of concomitant medications and/or medical histories
  - If relevant, summary of qualitative results
  - if relevant, statistical test results (including those from validation studies)
- References to animal or human toxicology, pharmacovigilance databases, participants or patient/consumer experiences (including explanations for differential reporting, such as qualitative results, and underlying conceptual theories or orientations)
- Key conclusions and limitations as reported by the original authors or as determined by the review author

Assessment of risk of bias in included studies

The risk of bias was independently assessed by EA and KB according to The Cochrane Collaboration's 'Risk of bias' tool, as far as was feasible in terms of the actual study design encountered (186). Where this was not feasible due to the study methodology (e.g. for reports that compared outcomes within-participants), the studies were critically evaluated

in terms of the potential impact of the study's design and conduct on its findings regarding selection, performance, detection, attrition, and reporting biases, and any other bias that the authors considered important.

It is acknowledged that a risk of bias assessment is dependent on the completeness and quality of the original study report and an attempt was made to contact study authors whose email addresses were available to retrieve protocols or specific relevant missing information (187). Reports were not excluded from this review on the basis of quality unless insufficient data for the comparison was reported.

Measures of the effect of the methods

Effect measures from different methods were compared where possible by assessing for an overlap in 95% confidence intervals (188).

Unit of analysis issues

The units of analysis were only known once the review was ongoing. The way that studies presented their data varied from absolute numbers of AEs, to means, medians, the proportion of participants with AEs and some study-specific scores. This precluded quantitative pooling of data to generate pooled estimates. However, wherever possible data were transformed into a common quantitative rubric.

Dealing with missing data

An aim was to minimize the amount of missing data through contact with authors as mentioned above (187). Thereafter, any assumptions made about missing data, any statistical methods used to impute them and the potential impact of these methods on the findings of the review, are reported.

Assessment of heterogeneity

As noted above, pooled estimates could not be calculated; therefore heterogeneity was not assessed as planned, using the Q test and I^2 statistic (189).

Assessment of reporting biases

As noted above, pooled estimates could not be calculated; therefore reporting bias could not be explored as planned using a funnel plot (190).

Data synthesis

As a meta-analysis of included studies was not possible given differences in study designs, interventions and presentation, a narrative synthesis was conducted by EA using recommendations by Popay et al. (191). Included studies were first examined for any *a priori* theoretical basis for how elicitation methods could differ, in case this could contribute to the interpretation and applicability of the review findings. A narrative summary of the scope of the study designs was developed in order to look at aspects of study design across studies, and the Characteristics of included studies table (Supplementary Table 3.1) was then used as the starting point for organising studies for synthesis. For the latter we tabulated brief key study characteristics and the results within two broad categories of whether outcomes were compared between- or within-participants. Where studies had not calculated an effect, this was performed using raw or summary data, where possible, in order to develop common quantitative rubrics. The tabulation was examined for relationships within and between the studies, the aim being to identify variables that potentially moderated the effects. The same process was followed, where relevant, for items relating to the impact of different elicitation methods on the between-drug effects, the nature of AEs detected and previous/concomitant medication and medical history data reported. The analysis was reviewed by the co-authors before finalisation.

Subgroup analysis and investigation of heterogeneity

A meta-analysis was not conducted therefore there were no quantitative subgroup analyses or investigations of heterogeneity.

Sensitivity analysis

A meta-analysis was not conducted therefore no sensitivity analyses were performed.

3.5 Results

3.5.1 Description of studies

See the Characteristics of included studies table (Supplementary Table 3.1), the Characteristics of excluded studies table (Supplementary Table 3.2) and the Characteristics of studies awaiting classification table (Supplementary Table 3.3).

Results of the search

See Figure 3.1 for a flow diagram of the search metrics and Supplementary File 3.1 for the electronic search results. The electronic search identified 13903 papers, decreasing to 9663 after de-duplication. An additional 35 papers were identified for inspection from non-electronic methods after reviewing the references lists of included and some relevant excluded reports and hand searches. 203 full texts in total were sought, of which 33 were eligible for inclusion. In total, 168 were excluded; 25 were further duplicates or variations on reports already assessed that did not add any relevant data, 10 were not in English or had been retracted, 98 did not include a comparison of methods for eliciting AE data, or the comparison was not possible due to the way data were collected or presented, 23 did not report any relevance for clinical trials, seven only included an objectively measured AE (e.g. observation by a healthcare worker or laboratory report), or otherwise ineligible assessment, and five were conference abstracts without an associated paper. See the Characteristics of excluded studies table (Supplementary Table 3.1) for details of those which readers may expect to be included. Nineteen included studies were found from the electronic search and 14 through non-electronic means.

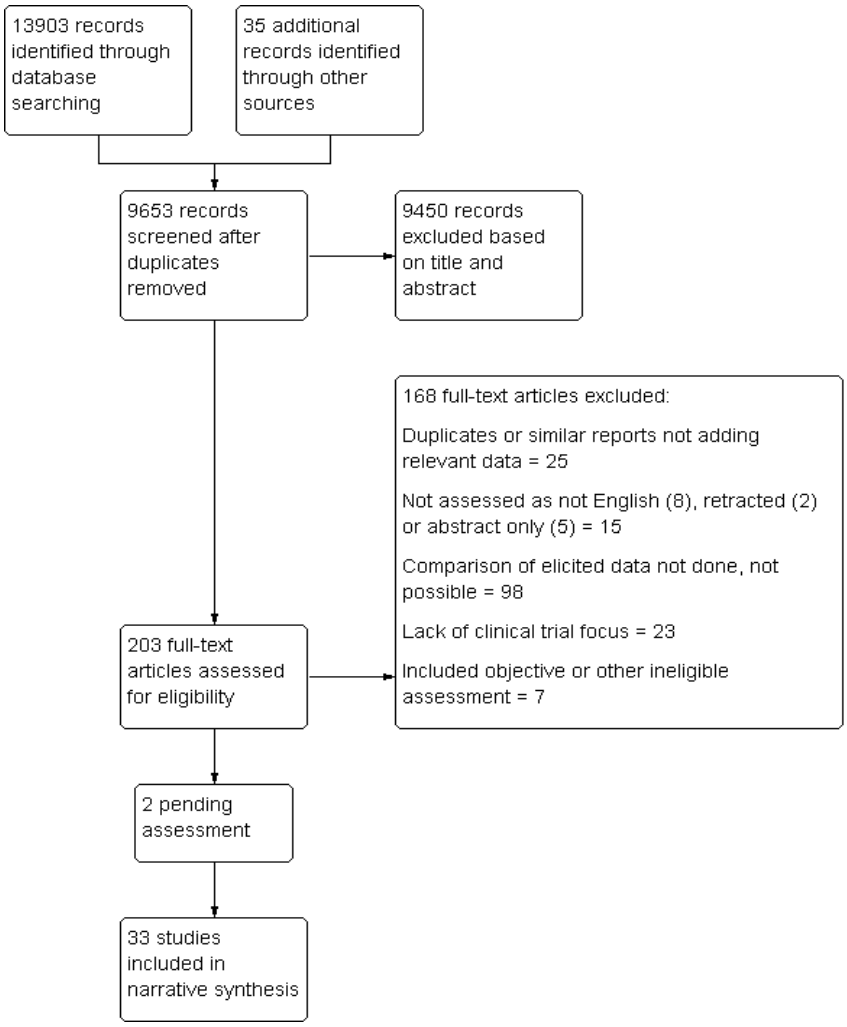

Figure 3.1: Flow diagram of the systematic review search metrics

Included studies

Of the 33 eligible studies, 32 were published in full and one, Kruft et al., as a letter to the editor (149).

3.5.2 Theoretical basis for elicitation methods

Studies were largely not explicit, or were unclear, about the theoretical basis for their work. Often the expectation was simply that data was likely to be underreported if participants are not questioned in detail. More specific questioning would be likely to increase the number of AE reports though potentially miss AEs not on the list/tool. Rabkin et al., however, asserted that a theoretical advantage of a general enquiry was the absence of suggestion, but also hypothesised that response styles (like stoicism) are more likely to influence open-ended questioning than checklist methods of elicitation (106). Avery et

al. postulated that patients allowed to volunteer information may conceal or fail to recognise symptoms, while suggestible patients may report symptoms when questioned where there is little objective evidence of the symptom being present (138). Rosenthal et al. stated that 'connotation of words, among many other factors' could influence AE responses (158). Some hypothesised that the nature of AEs detected by the different methods would be informative. This included the consideration that extra AEs reported through more specific questioning methods, like a checklist, were less likely to be clinically relevant, severe/bothersome or caused by the intervention than those reported spontaneously or in response to what some call 'non-leading' open enquiries (Barber 1995, Downing 1970, Rabkin 1992)(141, 146). Where hypotheses were mentioned they were largely based on a study or studies cited in this review that may or may not be supported by the evidence. Those that compared methods as part of a validation exercise for a new elicitation tool aimed to measure concordance (de Vries 2013 and de Vries 2014)(192, 193). Another study was explicit that it was conducted retrospectively with data already collected by different elicitation methods (Kruft 2007). Brent et al. and Monteiro et al. were natural experiments, in that the method for eliciting AE data changed during a trial (144, 152). For others it was unclear whether the comparison was an *a priori* objective (Hermans 1994; Nicholls 1980; O'Connell 2007; Os 1994; Wernicke 2005; Yeo 1991) (153-155, 164, 165, 194).

3.5.3 Scope of study designs and presentations

Methods

Included studies were conducted within a wide range of therapeutic areas (and therefore with various participants and drug interventions), including cardiology (Borghi 1984,Hermans 1994, Nicholls 1980, Os 1994, Reilly 1992, Rosenthal 1996, Török 1984, Wallander 1991, Yeo 1991) (143, 157, 161, 162), psychiatry (Avery 1967, Brent 2009, Downing 1970, Greenhill 2004, Jacobson 1987, Landén 2005, Monteiro 1987, Rabkin 1992)(112, 148, 150), ophthalmology (Barber 1995, Kruft 2007), diabetes (de Vries 2013, de Vries 2014), dysmenorrhoea (O'Connell 2007), gastro-intestinal diseases (Barrowman 1970)(142), gonorrhoea (Wallin 1981)(163), malaria (Allen 2013; the PhD candidate's own study reported in Chapter 5 of this thesis)(140), migraine (Sheftell 2004)(159), Parkinson's disease (Perez-Lloret 2012)(156), prostatic hyperplasia (Bent 2006)(139), rheumatology (Huskisson 1974)(147), and an unspecified indication for antihistamines (Lundberg 1980)(151). Three studies were not related to any specific therapeutic area (Ciccolunghi 1975, Spilker 1987) (145, 160) or the therapeutic area was not specified (Wernicke 2005).

Five studies involved healthy volunteers, students or employees (Allen 2013, Barrowman 1970, Ciccolunghi 1975, Lundberg 1980, Spilker 1987), the remainder were conducted with patients. Greenhill 2004 was treating children and therefore included caregivers in the elicitation process, some of whom received a reimbursement fee for taking part in the study (138). Two other studies were conducted in adolescents (Brent 2009, O'Connell 2007), one in children and adults (Wernicke 2005) and the remainder in adults (or assumed to be adults when age was not explicit; age ranges were in fact missing from several papers, so age was not included in the Characteristics of included studies table).
Most studies were conducted in Europe (N=17) or the United States (N=12). Two others were multinational (Kruft 2007, Sheftell 2004), one African (Allen 2013) and one conducted in an unknown location (Wernicke 2005).

Seven studies were conducted outside of a clinical trial (Ciccolunghi 1975, de Vries 2013, de Vries 2014, Greenhill 2004, Perez-Lloret 2012, Sheftell 2004, Spilker 1987), the remainder nested within, or integral to, a trial. The latter were all randomised clinical trials (RCT), with either a reference drug or placebo, except for Allen 2013 and Wallin 1981, who used single-arm designs, and Török 1984, who used a mixture of RCTs and single arm trials. Four studies were conducted as part of validating two new tools for eliciting AEs, de Vries 2013 and de Vries 2014 (outside of a trial) and Jacobson 1987 and Rabkin 1992 (within a trial).

Data

Most studies sought to elicit any AE, while others focused on a specific AE or specific AEs of special interest. The latter included occular-related abnormalities (Kruft 2007), sexual-dysfunction (Landén 2005, Monteiro 1987), depression (O'Connell 2007), cough (Os 1994, Yeo 1991), and self-harm (Brent 2009). However, there were significant variations in terminology and definitions for the data being collected, which made the analysis challenging. It was largely older studies that used the terms "side-effect" (Avery 1967, Huskisson 1974, Lundberg 1980, Nicholls 1980, O'Connell 2007, Os 1994, Török 1984), "side reaction" (Downing 1970), "unwanted effect" (Borghi 1984) and "adverse reaction" (Wallin 1981) to describe the data collected, regardless of whether it was treatment-emergent or not, or whether a causality assessment had been performed. More recent studies used "adverse experience" (Barber 1995), "adverse health event" (Jacobson 1987), "adverse event", "adverse drug event" or just "event" (Bent 2006, Brent 2009, de

Vries 2013, de Vries 2014, Greenhill 2004, Hermans 1994, Kruft 2007, Landén 2005, Perez-Lloret 2012, Rabkin 1992, Rosenthal 1996, Sheftell 2004, Wallander 1991, Wernicke 2005). As the protocol allowed for the inclusion of methods' studies conducted outside of a clinical trial it was expected that not all studies would collect and/or report AEs according to the ICH definition ('any untoward medical occurrence in a patient or clinical investigation subject administered a pharmaceutical product and that does not necessarily have a causal relationship with this treatment')(1). However, of the 26 comparisons nested within or integral to clinical trials, only six could confidently be considered as reporting treatment-emergent AEs - i.e. taking baseline symptoms or medical history into account when interpreting data collected as a new event or worsening of a previous event (Allen 2013, Brent 2009, Hermans 1994, Landén 2005, Monteiro 1987, Os 1994, Wernicke 2005). Four others were clear that they were collecting symptoms that were not necessarily treatment-emergent (Ciccolunghi 1975, Reilly 1992, Spilker 1987; and Török 1984). The remainder were unclear. AE is used as the default term hereafter, however, for simplicity. Rosenthal 1996 only reported on AEs that matched those on the CL.
Some studies looked at the nature of the AEs elicited, although the way this was done varied. Allen 2013 included a global statement of the investigator's assessment of severity and causality of all AEs collected. Other authors reported individual assessments of severity, clinical relevance, clinical action taken, seriousness, discomfort and/or bother. Assessments were either performed by investigators (Brent 2009, Greenhill 2004, Hermans 1994, Jacobson 1987, Perez-Lloret 2012, Rabkin 1992) or participants (Avery 1967, Barber 1995, Ciccolunghi 1975, Downing 1970, Reilly 1992, Sheftell 2004, Spilker 1987) using tools such as investigator grading schemes (mild, moderate, severe etc.) and participant-reported rating scales. It was not clear for Huskisson 1974 or Wallin 1981 who assessed severity.

Other relevant results related to the feasibility (de Vries 2013) and acceptability (Greenhill 2004) of different questioning methods, and de Vries 2014 measured differing recall periods. Allen 2013 collected qualitative data from in-depth interviews with selected participants to explore reasons for differential reporting between elicitation methods, and captured medical history and non-study drug data in the same way as the AEs. Avery 1967; Hermans 1994; Rabkin 1992 and Reilly 1992 also elicited baseline AE/symptom data, however there was not enough information to clarify whether measures of effect after baseline were treatment-emergent or not.

While studies that involved an objective assessment of AEs (e.g. through laboratory tests or physical examinations) were generally excluded, it was possible to identify some participant-reported AE data in two of these justifying their inclusion in this review (Török 1984, Wallander 1991). For another study (Borghi 1984) it is possible that one of the methods involved a doctor 'filtering' participant reports (i.e. just reporting those he/she considered ADRs). This study was also included in the review by restricting the comparison to the other two methods that reflected participant-reported data.

Comparisons

Most (25) studies compared data within participants. That is, each participant was asked about AEs by two or more elicitation methods. The remainder of the studies allocated groups of participants to different methods for eliciting AEs; Avery 1967, Bent 2006, Borghi 1984, Ciccolunghi 1975, and Spilker 1987 allocated methods randomly while Brent 2009, Huskisson 1974 and Török 1984 used non-random allocation. Four of the between-participant comparisons involved comparing one method with that method plus another one (Avery 1967, Huskisson 1974, Rabkin 1992 and Török 1984).

Most comparisons were of participant responses to an open question(s) (O) (and/or occasionally a completely spontaneous report, i.e. where no question was asked) and responses to what can be summarised as a checklist or questionnaire pick-list of potential AEs (CL). Two of these studies involved answers to an open question written on a blank (B) form Sheftell 2004, Spilker 1987) and one an open question on a daily diary (D) as a gold standard (de Vries 2014). Other comparisons involved rating scales (R), such as visual analogue scales (VAS), that used a particular change over time to determine the incidence of an AE (e.g. the Brief Suicide Severity Rating Scale used by Brent 2009) or was simply reported as a mm change (e.g. Os 1994). Two studies conducted interviews (INT) with participants (Allen 2013, Monteiro 1987). The comparisons other than O/B/D versus CL were: O versus R (Brent 2009, Kruft 2007, Landén 2005, Yeo 1991), CL versus another CL (de Vries 2013, specifically the impact of using body categories on reporting), CL versus R (Lundberg 1980, Wallander 1991), O versus CL versus another CL (Greenhill 2004), O versus CL versus R (Os 1994), O versus another type of O versus CL (Bent 2006), O versus CL versus INT (Allen 2013 [in a subset of participants only], Monteiro 1987), and B versus B versus CL (Ciccolunghi 1975).

The detail of questioning (e.g. phraseology of O, the number and type of specific symptoms or body systems asked about in CL), how methods were developed and applied (e.g. verbal, written or electronic) varied widely within these comparisons. See Supplementary Table 3.4 for more information.

Due to the variety of indications and interventions/treatments, and whether comparisons were within clinical trials or not, the timeline over which the elicitation methods were applied and how often they were applied were diverse; some compared data elicited on one visit occasion only, others on multiple visit occasions for which data were combined or reported separately.

Outcomes

In addition to differences in participants, therapeutic areas, comparisons, data and follow-up period, the disparate approaches to measuring and reporting outcomes supported our decision not to pool results.

3.5.4 Number of AEs reported

Several studies reported the number and/or proportion of AEs elicited by method at, or by, a particular time-point. These were either given as a sum total of all AEs, by type of AE or the raw data were listed by method. Three studies, however, only gave the number (%) of participants reporting no versus at least one or two AEs (Downing 1970,Landén 2005) or the mean/median (range, standard deviation) number of AEs per participant (Avery 1967, Downing 1970). Other variations included Barrowman 1970 andBarber 1995, who calculated an average frequency of a particular domain for participants who did not report any AEs spontaneously (O) but indicated an AE by CL, and Huskisson 1974, who presented AEs only as a score calculated from a severity rating scale. Kruft 2007 performed a meta-analysis of 4 studies where the outcome was presented as the number and proportion of participants reporting AEs. Os 1994 presented frequencies of AEs for two of the three methods compared, but presented the third method (R) as a change in VAS measurement. Three of the within-participant comparison studies only gave the number of additional AEs obtained through the second or third method (i.e. capturing AEs only when they were first elicited), rather than absolute numbers of AEs obtained by each method (Allen 2013, Greenhill 2004, Wallin 1981). Huskisson 1974 combined all AEs that were not auditory or gastro-intestinal in nature as 'irrelevant' and reported them as a combined frequency/severity score. Monteiro 1987 limited the comparison to participants who had

not reported an AE by CL but had reported at least one AE either spontaneously (O) or by INT.

Statistical tests of effect by elicitation method, where used, mostly included Chi-square and Mann-Whitney U tests. However, de Vries 2013 also used normal curve deviate statistics (Z value) for the measure of agreement between methods, while de Vries 2014 calculated the sensitivity and positive predictive value at different MedDRA® reporting levels. Wallander 1991 investigated the ability of each method to detect symptoms that change over time and sensitivity to change, while Perez-Lloret 2012 analysed factors related to spontaneous reporting of AEs.

Seventeen studies presented comparative data by trial drug arm, either descriptively or through measures of effect (Avery 1967, Barrowman 1970, Borghi 1984, Ciccolunghi 1975, Downing 1970, Hermans 1994, Huskisson 1974, Landén 2005, Lundberg 1980, Monteiro 1987, Nicholls 1980, O'Connell 2007, Os 1994, Reilly 1992, Rosenthal 1996, Wallander 1991, Wernicke 2005). However, the data in Monteiro 1987 could not be extracted for both arms. Wernicke 2005 used the ratio between the rate of AEs reported by drug versus placebo (D/P) plotted for solicited AEs on an x-axis against spontaneous AEs on a y-axis, and the ratio of D/P ratios. Lundberg 1980 used an analysis of variance for data in each method arm reported by at least 50% of the sample, however only summary descriptions were available for the between-method comparison results. Landén 2005 and Os 1994 also reported AEs by drug and gender for elicitation methods compared.

3.5.3 Nature of AEs reported, including quality of life, clinical relevance and action taken

The way studies analysed and presented the effect of elicitation method on the nature of AEs reported by participants also varied. Avery 1967 compared mean severity score on each visit occasion (weighting symptoms by a factor derived from the degree of subjective discomfort reported by the patient). Barber 1995 reported the average bother, level of activity limitation, satisfaction with medication, compliance, and global quality of life domain scores for those patients who did not report any spontaneous AEs (O) but did indicate on the questionnaire (CL) that they had the AE. Downing 1970 compared the levels of intensity of AEs identified by CL but not O, with those detected by both methods. Ciccolunghi 1975 reported the frequency and proportion of AEs by discomfort level, and by elicitation method, Greenhill 2004 and Hermans 1994reported the number

(%) of AEs by severity. Huskisson 1974 assigned a score based on whether an AE was absent, slight, moderate or severe, thus combining frequency and severity as the only outcome. Jacobson 1987 and Rabkin 1992 reported mean AE severity and impairment by elicitation method. Reilly 1992 presented the proportion bothered, the mean degree of bother per patient, and calculated a severity index score. Rabkin 1992 and Greenhill 2004 both considered the impact of elicitation method on the number of participants for whom clinical action was taken and the number (%) of AEs deemed clinically relevant. Two studies looked at the duration of AEs detected by different elicitation methods (Hermans 1994, Reilly 1992) while Brent 2009 measured time to onset of self-harm AEs by elicitation method. Ciccolunghi 1975 and Spilker 1987 described the most commonly reported AEs by elicitation method, while Jacobson 1987 and Rabkin 1992investigated categories of AEs that were reported by participants through use of different questioning methods; Rabkin 1992 specifically looked at whether certain AEs are underreported through O (e.g. sexual dysfunction) by highlighting AEs that were reported more than five times as often by CL compared to O.

3.5.6 Other relevant outcome variables

Other outcomes reported by some authors related to tool validation; de Vries 2013 investigated content validity through cognitive debriefing, while de Vries 2014 compared four-week versus three-month recall periods and Jacobson 1987 investigated interrater reliability between different trial staff. de Vries 2013 also examined the feasibility and acceptability of the different elicitation methods by time to complete the CLs and asking about ease of use. Greenhill 2004 determined the proportion of clinicians and parents rating satisfaction with the elicitation methods using several domains.

Allen 2013 performed a thematic qualitative data analysis in terms of explanations given for differential reporting of AEs, medical history and the use of non-study medications, and how participants expressed themselves, exploring the emerging themes in relation to broader theories.

Excluded studies

As the search was broad out of necessity, only those studies which a reader may reasonably expect to see among the included studies are listed in the Characteristics of excluded studies table (Supplementary Table 3.2). As indicted in the results of the

search above, studies were excluded if they were not relevant to the topic of the review (e.g. non-drug studies), if only abstracts were available where there was not enough information for an assessment or inclusion (and no relevant paper found), if there was no relevance to clinical trials, if there was no comparison of elicitation methods or data were presented in a way that could be extracted. For the latter, some studies at first glance appeared to make a comparison but had to be excluded as the data collected through the different elicitation methods were categorised differently.

3.5.7 Risk of bias in included studies

See Supplementary Tables 3.5a and b.

Allocation

For all 25 within-participant comparison studies there was a technically high risk of bias as neither random sequence generation and allocation concealment is feasible to implement. However, these are unlikely to have impacted the results substantially as all participants were exposed to all methods. For two studies groups the order that participants received the questions were also randomly assigned (de Vries 2013, Jacobson 1987). For the eight between-participant comparisons, Bent 2006 used a computer-generated method and blinded personnel to the allocation, therefore had a low risk of bias. Ciccolunghi 1975 used a pre-determined allocation list and Spilker 1987 used a table of random numbers which also reduced the risk of bias. The remaining five studies were either deemed high risk of bias as it was clear the allocation was not random (Brent 2009, Huskisson 1974, Török 1984), or the allocation method was unclear (Avery 1967, Borghi 1984). However, where groups were determined by site (Huskisson 1974), it may be that there is a lower risk of bias as different staff are involved. The latter, in itself, however raises the possibility for inconsistent recording of AEs. Ciccolunghi 1975 had a large proportion of non-responders to the invitation to take part in their study which may increase the risk of bias as there may have been a selective non-response related to a particular question method.

Blinding

Of the 33 studies included in this review, only one stated that the investigator was blinded to the AE data reports compared in their studies; Borghi 1984 reported that the investigator was neither informed of the results of the self-reported AEs by either the self-completed checklist or blank form, nor did he/she help participants fill in the forms. It is not feasible to blind participants when they are reporting AEs, however for those taking part in

studies where the comparison was between-participants there may be less risk of bias from knowing which questioning method was used, as it is less likely that participants were made aware of different methods being used in different groups/sites. For the within-participant comparisons it is highly likely that participants were 'primed' by the first method - i.e. they would be more likely to report that same AE when the second method was applied. As such, these are not independent comparisons, which was acknowledged by some study authors (Allen 2013, Greenhill 2004).The risk of bias for Wallander 1991 may have been the highest as participants took forms for both questioning methods home to complete on their own (there were no details in the paper as to instructions for their order of completion so one may prompt what to report in the other).

Follow up and exclusions

Most studies had a low risk of attrition bias. However this was unclear in nine studies and two were considered of higher risk as they had high dropout rates and it could not be sure that these were not related to the questioning method (Ciccolunghi 1975, Wallander 1991). For Wallander 1991 there was a significant number of dropouts, which could potentially relate to a questioning methods as participants took forms for both questioning method home to complete; they may have decided not to bring one of the forms back due to the nature of its questions. For Ciccolunghi 1975 there was also a significant amount of missing data, which could potentially be related to the method of questioning as the forms were distributed by internal mail and it was therefore up to staff whether they completed the forms or not. It is possible that the decision not to complete and return a form was affected by the type of questioning (e.g. a longer form less likely to be completed that a shorter one)

Selective reporting

Twenty studies had a low risk of selective reporting and in 11 it was unclear. Two studies were deemed high risk as they only presented a selection of data; Rosenthal 1996only presenting spontaneously-elicited AEs if they matched questions on the questionnaire that they were being compared to, and Wernicke 2005 selecting for comparison those AEs reflecting the same symptom in the spontaneous and solicited methods (in order to calculate ratios of the rate reported by drug versus placebo). While there were clear reasons for these selections of data, it is possible that data not selected may have been informative.

Other potential sources of bias

For Barber 1995 it was not clear whether the elicitation methods were applied in the same order for all participants, while for Os 1994 it was unclear as to how questioning methods were applied, therefore it was not possible to assess whether the application could have biased these studies in some way. The application of elicitation methods was unclear for Borghi 1984, Huskisson 1974, Monteiro 1987, O'Connell 2007, Nicholls 1980, Reilly 1992, Rosenthal 1996 and Török 1984. In addition, there may have been differences in the phrasing of open-questioning in Allen 2013 and Barber 1995. For the between-participant comparisons used in Brent 2009 and Huskisson 1974, it was not clear whether one group of participants were exposed to both questioning methods. Kruft 2007 and Wernicke 2005 were meta-analyses where there was little information about the parent studies so it is unclear if there may have been some other inherent biases. Finally, all studies are limited by a lack of a true gold standard against which to assess the data reported by participants.

3.5.8 Effects of methods

3.5.8.1 The impact of different elicitation methods on the number of AEs reported

Between-participant comparisons

- Overall

For the eight studies comparing elicitation methods between groups of participants, 12 different comparisons involving an open enquiry (O) could be derived (one of which had three different endpoints, resulting in 14 comparisons in total) (Supplementary Table 3.6a). There was no common statistical rubric; however it was possible to re-present some effect measures as a risk ratio of the number (%) of participants with at least one AE. Using O as the reference, there was an increase in the absolute or mean number of AEs elicited, or the number of participants reporting at least one AE, whenever CL or R was used (except for suicide attempts in Brent 2009 which were more often reported by O than CL). This increased sensitivity of CL/R was observed regardless of the study location, therapy area, whether the study was conducted within or outside of a clinical trial or with patient or healthy volunteers, the duration of follow-up and whether the outcome variable could be considered as a treatment-emergent AE or not. The two studies that compared different types of O (Bent 2006 and Ciccolunghi 1975) found no difference in the number of AEs detected.

- Between study arms

Four studies presented data by study arm (Avery 1967, Borghi 1984, and Huskisson 1974 within a clinical trial and Ciccolunghi 1975 outside of a clinical trial). Avery 1967 found that the trend for more AEs through CL was sustained when the active arm was examined while removing the placebo participants. Borghi 1984 found that there did not appear to be a difference between the methods for detecting AEs. Huskisson 1974 had pre-determined that only auditory and GI AEs would be termed drug-related and all other AEs were deemed as irrelevant "noise". Using this classification system they showed that the reporting of all three types of AEs increased for fenoprofen by two to three times, although the aspirin scores were inconsistent. While a different context, for Ciccolunghi 1975, there was no difference in the detection of AEs between those participants taking medication and those not taking medication when comparing the two types of O. However, there was a statistically significantly greater number of AEs detected by CL compared to O in those participants taking medications, but no such difference seen in those not taking medications.

Within-participant comparisons

- Overall

For the 25 studies comparing elicitation methods between groups of participants, 19 comparisons involved an open enquiry (O/B/D) and a checklist-type method (CL) (although for Kruft 2007 the data for CL and R could not be distinguished). See Supplementary Table 3.6b for a summary of the results. The direction of the effect of the method on the number of AEs was in favour of the CL in all cases except for Hermans 1994 and de Vries 2014. The former, despite finding an increase in absolute numbers of AEs with the CL, found no increase when looking at the percentage of participants with at least one AE. de Vries 2014 had performed their study to validate a CL and found low sensitivity (33%) and positive prediction values (10 to 51%) compared to their open question diary. The fact that the diary was completed daily will surely have influenced this finding.

Two studies compared different types of CL - de Vries 2013 and Greenhill 2004. The former found that adding body categories did not affect the frequency of AE data reports, while Greenhill 2004 found that using a body system review resulted in a greater increase in AE reports compared to a drug-specific inquiry.

Several studies incorporated scales (R); Landén 2005, Kruft 2007, Yeo 1991 and Wallander 1991 found that the use of R resulted in increased AE reports. Os 1994 observed that the

increase in cough reported by R was less consistent in men compared to women. Landén 2005 found less women than men reporting AEs by O but more women than men reporting AEs by R (the latter not statistically significant).

Monteiro 1987 found 36% of those with drug-induced sexual-dysfunction at INT did not report this at the previous CL, despite their concern about this AE, and even if they were secretly reducing dose of drug to overcome it.

Perez-Lloret 2012 explored the relationship of various demographic and disease related factors with reporting at least one AE in response to O, and the only association found was with participants who reported more than two AEs by CL.

- Between study arms

Common findings were identified when considering if the question method influences the ability to detect differences between arms. Nicholls 1980 and Rosenthal 1996(drug-drug comparisons) and Downing 1970 (drug-placebo comparisons) showed a statistically significant difference between trial arms when using CL, and no such effect when using O. Hermans 1994 found the opposite; the between-drug difference in AEs overall (and for frequency of ankle oedema) was statistically significant for O, not for CL. Wernicke 2005's use of drug/placebo ratios for reported AEs suggested that O is more effective in distinguishing a difference between trial arms . However, they also found that there were more statistically significant differences between trial arms by CL compared to O (nine versus five AEs). Landén 2005 (drug-drug) and O'Connell 2007(drug-placebo) showed no difference between arms. The two studies that compared a CL with a R also had conflicting results - Lundberg 1980 finding a difference between drugs by their CL but not by their R, while Os 1994 found no difference between these types of tools.

3.5.8.2 The impact of different elicitation methods on the nature of AEs reported

Between-participant comparisons

Avery 1967 found a statistically significant higher mean severity at each visit by CL both overall and when removing placebo data, in contrast to Ciccolunghi 1975 who found that O was associated with a greater severity of symptoms than CL (Supplementary Table 3.6c). Brent 2009 found no difference between O and R for reporting of serious suicidal or non-suicidal ideation AEs, but the time to onset for both was earlier for data elicited by R compared to O. In terms of the individual types of AEs reported, Spilker 1987 found that

the most common symptoms by CL were fatigue, headache and nasal congestion and by O were headache, back or muscle pain and nasal congestion, so there appears to be some overlap. Huskisson 1974 had used a composite measure for frequency and severity so is reported under the number of AEs section above.

Within-participant comparisons

Of the 10 studies investigating the nature of AEs reported through questioning method, six found O more likely to detect more severe or intense AEs, or AEs causing more bother, distress or limiting activity (Barber 1995, Downing 1970; Greenhill 2004, Jacobson 1987; Rabkin 1992; Reilly 1992), though Greenhill 2004 showed that their drug-specific review CL detected more moderate AEs compared to O (Supplementary Table 3.6d). Meanwhile, the paper by Rabkin 1992 revealed that 61% of AEs rated severe or very severe were elicited by the CL and 65% of AEs causing severe or very severe dysfunction were detected by CL compared to 35% by O. Allen 2013 reported that additional AEs detected thorough CL or INT were rated as mild, however the severity of AEs detected by O was not given. Hermans 1994 and Perez-Lloret 2012 found no difference between the elicitation methods in the severity of AEs detected, as did Sheftell 2004. However the latter also found that 31 (7.5%) of participants who rated their AE as severe in the CL had not reported it when previously asked by O.

Barber 1995 found that patients who spontaneously reported AEs indicated a more negative impact of side-effects and activity imitations on quality of life, more dissatisfaction with their medication, and more non-compliance compared to those not reporting spontaneously; the average global quality of life scores increased as patients reported AEs spontaneously and discontinued therapy. The two studies that looked at the duration of AEs detected by questioning method (Hermans 1994 and Reilly 1992) showed no difference. Jacobson 1987 observed that their CL detected a greater variety of AEs compared to O. Rabkin 1992, using essentially the same tool, observed that the 23 AEs that were reported more than five times as often by CL compared to O included no reports of sexual dysfunction. The authors conclude that there was therefore no evidence of selective underreporting of sexual dysfunction by O compared to other AEs such as cognitive and affective symptoms. They suggest that the phrase used to introduce the questioning may have inadvertently suggested to participants that the latter were not the topic of the enquiry.

3.5.8.3 The impact of different elicitation methods on other relevant outcome variables

Clinical action/relevance

Rabkin 1992 found that both their O and CL methods contributed equally to the elicitation of AEs that, in the clinician's opinion, required some change in management (13% and 11% respectively). However, those AEs elicited by O required more extensive changes (dose suspension or discontinuation compared with increased surveillance or change in dose) and all of these AEs were in the trials' drug arms, not placebo. Greenhill 2004 also found a higher proportion of AEs elicited through O led to some clinical action (31%) compared with those identified with the CLs (12% when using a drug-specific CL and 15% a body system review). Of the clinically relevant AEs (n=37), 19 (53%) were elicited by a body-system review.

Validity, feasibility, acceptability and satisfaction

de Vries 2013 found significant problems during content validation that needed to be resolved while designing their self-reported questionnaire (CL), such as making it clearer to participants which questions related to their underlying disease symptoms and which to AEs. For the final tool, a questionnaire of 252 items, approximately 50% of respondents found a body category structure to be helpful, while most found the tool easy to use and took less than 60 minutes to complete. Greenhill 2004, meanwhile, found that while 80% of parents found their body system review CL method "just right" (71% specifically finding the duration of the process "just right"), 70% of clinicians considered it to be too detailed and 74% found it took too long. Of note, however, is that satisfaction ratings for the more detailed enquiry were significantly higher in parents who received reimbursement fees compared to those who did not. Overall the body system review was deemed very useful by 66% of parents but only 28% of clinicians in that study. de Vries 2014 found that a recall period of either four weeks or three months did not impact the sensitivity of identifying patients who experienced an AE through their CL tool at either a MedDRA® organ class or specific level. The positive predictor value was especially low for a four-week recall period. Jacobson 1987found good overall interrater reliability for detecting AEs using both their O and CL methods (best when raters were both present for the same patient consultation) but low interrater reliability for individual AEs and measures such as duration, severity and functional impairment.

Qualitative

The one study that incorporated a qualitative analysis (Allen 2013) found that their CL and INT facilitated participants' recognition of health issues and treatments, and consideration of what to report. Information about AEs, medical history and/or non-study medicine use was sometimes not reported because participants forgot, it was considered irrelevant or insignificant, or they feared the consequences of reporting. Some medicine names were not known, and answers to questions were sometimes considered inferior to what blood tests can do for detecting ill health. There were some differences between the two trial sites in this study that had an impact on reporting - South African inpatient HIV infected but otherwise healthy volunteers exhibited a "trial citizenship", working to achieve the researchers' goals, while Tanzanian HIV positive or negative outpatients with malaria symptoms sometimes deferred responsibility for identifying items to report to the trial's clinicians.

Non-study medicines and medical histories

Allen 2013 also found that using the CL and INT (the latter in a subset of participants) after O resulted in an additional 23 and four non-study medication reports respectively in one site, two and nine in the other. The same pattern was found for past medical history reports; an additional eight and four reports for CL and INT respectively in one site, 245 and 15 in the other site. These quantitative data could not, however, be pooled due to different numbers and types of participants in each site.

3.6 Discussion

This review shows that the question of how different elicitation methods impact the reporting of subjective AE data by drug trial participants has been considered since at least the 1960s, and yet is still being debated nearly 50 years later. This situation probably reflects the complexity of the topic - how to accurately represent the often unknown adverse effects of a drug on a myriad subjective end points. The review itself was complicated by the diversity in included studies' participant populations, designs and elicitation methods. It is also difficult to ensure quality in this kind of methodology research and how to differentiate AE reports from disease-related symptoms. However the review does provide reasonable evidence of an increase, often substantial, in the number of AEs elicited when using more comprehensive (specific, detailed and/or lengthy) questioning, whether a checklist-type tool or rating scale, compared to a more open general enquiry,

whether a verbal question or blank form for self-report, in a wide variety of indications and contexts. This finding is, of course, intuitive.

Importantly, some of the included studies took their research beyond the quantitative effects and investigated the nature of the AEs reported in response to different questioning methods. Of the 10 studies comparing elicitation methods within-participants, six found that participants reported more severe/bothersome AEs to an open enquiry. While the studies used disparate methods to assess these end points (including the detail of the questioning itself and whether the nature of the AE was assessed by participants or investigators), these findings are supported by the qualitative data from Allen 2013, whereby participants described this process in action; the checklist had reminded them of a mild or intermittent AE, or of the need to consider and/or report it. However, some studies had different findings - Rabkin 1992 showed that even quite severe AEs were missed by the open enquiry and only detected by a checklist. Moreover, Monteiro 1987 found that some debilitating sexual dysfunction AEs were not reported spontaneously or in response to a specific checklist and only were revealed at an in-depth interview. Rabkin 1992 also suggested that the way the instructions to participants had been phrased in their general enquiry may have resulted in under-reporting of cognitive and affective AEs. This signals the care required when phrasing questions. It is therefore difficult to draw firm conclusions about the impact of questioning method on the nature of AEs detected without further work.

The work that was nested within comparative drug trials also had mixed results when considering if the question method influences the ability to detect differences in harm between trial arms. Nicholls 1980 and Rosenthal 1996 (drug-drug comparisons) and Downing 1970 (drug-placebo comparisons) showed a statistically significant difference between arms when using a checklist-type tool, and no such effect when using an open enquiry. Meanwhile Hermans 1994 found the opposite; the open enquiry appeared to detect a difference while the checklist did not. Wernicke 2005 used a drug/placebo ratio for reported AE, which suggested that the open enquiry is more effective in distinguishing a difference between arms. However, they also found that there were more statistically significant differences between trial arms by a checklist approach compared to an open enquiry (nine versus five AEs). Borghi 1984 , Landén 2005 (drug-drug) and O'Connell 2007 (drug-placebo) all showed no difference between question methods, so no conclusions could be drawn for those studies comparing checklist-type tools with rating

scales. These mixed results may reflect several issues: problems with the parent trials relating to their power to detect a difference in the safety outcome by drug, that a particular questioning method is no better than another at being able to detect a difference between arms, or that the arms being compared have similar safety profiles. One response to this review's findings would be to suggest that all studies use comprehensive specific enquiries in addition to an open enquiry, as they appear to be complementary. This would fit with the tenet of many clinical trials, certainty pre-registration trials, which aim to collect all AEs (i.e. high sensitivity with no provision for specificity relating to AE severity, bother from AEs, the clinical outcome or association between the drug and the AE). However more comprehensive enquiries are time-consuming and, while Greenhill 2004 found that parents of children generally found the more detailed questioning useful, a majority of clinicians found using a body system review too detailed and took too long. It is also unclear as to how much questioning is comprehensive enough, bearing in mind that tools may range from a short checklist to up to the 252 items developed by de Vries 2013. More work is therefore needed to explore the practicalities of using tools of different lengths and designs and achieve a balance between sensitivity and feasibility. For instance de Vries 2014 found that adding body categories which filtered and therefore limited the questions did not affect the outcome.

Another option is to use different types of questioning depending on what is known about the safety profile of the drug, which may itself be a factor of the phase of development; more comprehensive questioning early in the plan and less comprehensive questioning as data builds about a favourable benefit:harm profile. Provision for this is made by the US FDA - manufacturers are allowed to make a case for limiting the safety data to serious AEs for instance (91). In addition, for the treatment of life-threatening illness like severe malaria, there may be a case for using less a sensitive enquiry and focus on the more serious AEs. However, while it is unlikely that there is a perfect questioning tool, it is difficult to recommend that researchers use an enquiry method that may miss some clinically relevant data or effects important to patients, especially those that could impact on adherence when a drug is eventually distributed on a large scale post-registration. Despite what researchers and regulatory authorities may feel is known about important ADRs, there may be long-term or persistent mild effects that do not influence clinical action but nevertheless impact on the quality of life of even severely ill patients. An example is persistent nausea in oncology patients that is considered less clinically relevant by clinicians but a debilitating disorder by patients (48). Similarly, while there is guidance about the

need to enquire specifically for potential ADRs that are embarrassing for participants to talk about (such as sexual dysfunction), other drugs with the potential for such affects may not been identified as yet and this ADR will only be detected after many patients have been exposed to the drug (4).

The findings of this review are pertinent considerations for any clinical trial, however notably there was a dearth of relevant studies conducted in resource-poor settings. As mentioned in Chapter 2, this informed the rationale for designing and then including in this review the study reported in Chapter 5. As described in Chapter 1, the collection of data in these regions may be more difficult to achieve but essential as their populations, including vulnerable groups such as young children and pregnant women or those with significant comorbidities, are being widely exposed to new drugs such as the antimalarials. It is therefore critical to know which methods are used to question participants taking part in antimalarial clinical trials, and other pharmacoepidemiology studies, about their health and use of treatments in order to understand the potential impact of the method(s) on the safety data elicited. This in turn will help interpretation of individual and pooled safety analyses which form the basis of treatment policies in these regions with a high burden of diseases like malaria. It will be important to conduct further methodological research with the range of populations taking part in such research, including adults, young children and pregnant women, in order to investigate more closely any specific influential factors.

3.7 Summary of main results

Despite different study designs, populations and details of questioning methods, the review showed that more specific questioning of study participants leads to more AEs being reported compared to a more general enquiry. A subset of six studies suggests that more severe, bothersome or otherwise clinically relevant AEs were reported when an initial open enquiry was used while some less severe, bothersome or clinically relevant AEs were only reported with a subsequent specific enquiry. However one study showed that quite severe or debilitating AEs were only detected by an interview, while others did not find a difference in the nature of AEs between elicitation methods. No conclusions could be made regarding the impact of question method on the ability to detect a statistically significant difference between study arms as the findings were inconsistent.

3.8 Overall completeness and applicability of evidence

This review shows in a wide variety of populations that more AEs will be reported when participants are asked more comprehensive questions about their health. Some authors of studies included in this review hypothesised that such intensive questioning is suggestive, the implication being that participants will be made to report an AE that is not actually real. This review did not uncover any evidence for these concerns as there was no gold standard applied in the studies against which to measure the "truth", although a diary may be the closest to this (192). More work could be done to harness the techniques used in patient-reported outcome (PRO) methods to understand and validate AE context-specific question tools as is being done for oncology (119). As a minimum there is a need for more studies using interviews to understand reporting behaviour in a variety of contexts.

Another observation from this review's findings is that there were few tools other than checklists and scales used, aside from one diary. This could be an issue of the review methodology and studies comparing such tools may have been missed. Alternatively, it could be that studies use such tools but do not perform methodology research, or that they are seldom used to elicit AEs. While there has been significant technological growth in innovative ways to engage with trial participants about a range of experiences and end points, including health-related quality of life measures, the methods for eliciting AEs are somewhat lagging behind. Until there is progress authors should be encouraged to be clear within their teams as to the rationale for, and application of, the chosen questioning method used, as this will contribute to consistency in trial conduct. As important is the need to provide sufficient detail of the elicitation method when reporting results as this will lead to a better understanding by readers about how this may have influenced individual results and help in the conduct of systematic reviews and meta-analyses.

Few studies considered the impact of the elicitation method on other variables such as previous or concomitant medications and medical histories. These are important for determining whether AEs are ADRs (and may impact on eligibility criteria). This finding may be because the primary outcome of this review related to AEs and did not identify other studies that focused exclusively on those other variables.

3.9 Quality of the evidence

The Grading Recommendations Assessment, Development and Evaluation (GRADE) approach was not used in this review as there was no meta-analysis possible. The quality of evidence is limited by the heterogeneity of studies included, the design limitations inherent for this kind of methodological research but also limitations in the application of study methods and incomplete reporting in individual studies. Aside from differences in therapeutic areas, drug interventions and the actual questioning methods applied, some studies were conducted with students or staff (Ciccolunghi 1975 and Spilker 1987) as opposed to trial participants or patients. While there was merit in being inclusive in anticipation that the number of eligible studies may be low, these differences do limit the ability to make recommendations about specific elicitation methods or contexts. Sequence generation and allocation concealment are not necessarily relevant for within-participant comparison studies but are for between-participant comparisons, particularly for studies with non-random allocation of elicitation method to participants enrolled at the same site (Avery 1967, Brent 2009). Where elicitation methods were allocated by site or study (as for Huskisson 1974 and Török 1984) it is more likely that there was inconsistent recording of AEs, which could also impact the quality of the comparisons. It is not feasible to blind participants to a questioning method and this is likely to have affected the reporting of AEs to consecutive elicitation methods - the data reports are not independent. There may have been a particular risk of bias for Wallander 1991 as participants took forms for both questioning methods home to complete on their own. However for the other studies, if the data were applied consistently and recorded accurately then these cumulative comparisons are still useful, as they reflect the impact of using more than one method together, compared to one of the methods being used alone. All studies are limited by a lack of a true gold standard against which to assess the data reported by participant. Most studies had a low or unclear risk of attrition bias. However two can be considered of higher risk as it could not be sure that the drop-outs were not related to the questioning method used in those studies (Ciccolunghi 1975, Wallander 1991).

3.10 Potential biases in the review process

There were practical reasons for limiting the search to English studies reporting terms synonymous with AE three or more times in the title or abstract. The review could, however, have been improved by extending the search to languages other than English. In particular studies conducted and reported in French and Portuguese would be informative

for research conducted in francophone and lusophone Africa and may, in particular, have uncovered relevant work conducted in antimalarial clinical trials. The results also suggest that the electronic search missed a significant number of publications identified by non-electronic means. This highlights the importance of conducting a thorough review of reference lists of both included studies and excluded studies that nevertheless are relevant to the topic, as was done for this review. This finding, however, does raise the possibility that other eligible studies may have been missed, with an adverse impact on the overall conclusions of the review. It was considered unlikely that changes to the search strategy could be improved without increasing the number of references to an unmanageable level. However, this issue should be explored in further methodological research.

While studies that included objective measures of AEs were excluded unless the subjective data could be extracted separately, some included studies may have also included objectively measured AEs without reporting them as such. The review included populations taking part in clinical trials and methods studies outside of a clinical trials provided relevance to clinical trials was cited. These environments may be quite different in that the trial context may shape behaviour, including how AEs are reported (140, 195-197). Separating the three-way comparisons of within-participant studies into two-way comparisons may have distorted results due the possible effects of priming of one method on a subsequent one. The review may have been adversely affected by limiting the search strategy to English, as potentially relevant papers in other languages may have been missed. We chose to include studies regardless of whether they prospectively or retrospectively addressed the comparison of AE elicitation methods (and this was not always stated in the papers), although in general it is less optimal to retrospectively report outcomes. Lastly, it was not always possible to confirm the authors' calculations.

3.11 Agreements and disagreements with other studies or reviews

No similar studies were identified.

3.12 Authors' conclusions

3.12.1 Implications for systematic reviews and evaluations of healthcare

This review supports concerns that the methods used to elicit participant-reported AEs influence the detection of these data. There is the risk for under-detection of AEs, some of which may be important from a clinical perspective or for patients themselves, in studies

using a general elicitation method compared to those using a more comprehensive elicitation method. This could compromise the ability to compare the results from studies or pool data. However the impact on the nature of the AE detected by different methods is unclear. Regulatory authorities and others who guide researchers should emphasise how different elicitation methods have this potential to impact on study results and encourage drug research communities to debate working towards harmonisation of how safety data are collected, assessed and reported in their field. However, any chosen questioning method need to be feasible to use for both staff and participants. These issues are important for all clinical trials, including those investigating the effects of antimalarial drugs in large, often vulnerable, populations.

3.12.2 Implications for methodological research

The wide variety and inherent low quality of methods used to compare elicitation strategies in clinical drug trials limited this review. Though a complicated area, future studies would be improved by using and reporting clear definitions and terminology for AEs (and other important variables), the frequency and time period over which they were ascertained, and how they were graded, assessed for a relationship to the study drug, coded and tabulated/reported. As for similar work conducted in other areas of pharmacoepidemiology, this work is hampered by the lack of a true gold standard against which to assess the data reported by participants; measures are likely to be of concordance rather than validity (173). However, improving record linkages, using blood samples for pharmacokinetic analysis of commonly used medications to detect non-study medicines, and comparing new strategies against existing ones are options to explore. While the many potential AE end points in a given trial may preclude the development of general AE PRO measurement instruments, much could also be learnt from how these employ both quantitative and qualitative methods to understand data elicited (7). There was a dearth of relevant studies involving populations in resource-poor settings, including antimalarial studies and those involving pregnant women. It will therefore be important to conduct relevant methodological research with such populations that carry the highest malaria burden in order to investigate more closely any similarities and differences between how questioning methods influence safety data end points.

3.13 Acknowledgements

The review authors would like to acknowledge the Cochrane Adverse Effects Methods Group for their peer review of the protocol. We would also like to thank Mike Clarke and Clare Chandler for reviewing aspects of the protocol during development, the trials search coordinator Fiona Beyer, and Tamara Kredo for her comments during the data analysis.

3.14 Declarations of interest

The review authors have no interests to declare.

3.15 Differences between protocol and review

The review was clarified as regards the definition of AEs in 'Types of data", effect measures in 'Types of outcome measures', that searches were limited to English and the use of Popay 2006 to guide the narrative synthesis.

CHAPTER 4: EVALUATING HARM ASSOCIATED WITH ANTI-MALARIAL DRUGS; A SURVEY OF METHODS USED BY CLINICAL RESEARCHERS TO ELICIT, ASSESS AND RECORD PARTICIPANT-REPORTED ADVERSE EVENTS AND RELATED DATA

Elizabeth Allen[1], Clare Chandler[2], Nyaradzo Mandimika[1], Cheryl Pace[3], Ushma Mehta[1], Karen I Barnes[1]

[1] Division of Clinical Pharmacology, Department of Medicine, University of Cape Town, Cape Town, South Africa.
[2] Department of Global Health & Development, London School of Hygiene & Tropical Medicine, London, United Kingdom
[3] Department of Clinical Sciences, Liverpool School of Tropical Medicine, Liverpool, United Kingdom

Publication status:

Published: Malaria Journal 2013;12:325.

Author contributions

The candidate (EA) conceived of the survey, led design, data collection and analysis, and wrote the paper. CC, UM, CP and KB had input throughout while NM conducted data management and contributed to data analysis. All authors approved the final paper.

Abstract

Background: Participant reports of medical histories, adverse events (AE) and non-study drugs are integral to evaluating harm in clinical research. However, interpreting or synthesizing results is complicated if studies use different methods for ascertaining and assessing these data. To explore how these data are obtained in malaria drug studies, a descriptive online survey of clinical researchers was conducted during 2012 and 2013.

Methods: The survey was advertised through e-mails, collaborators and at conferences. Questions aimed to capture the detail, rationale and application of methods used to obtain relevant data within various study designs and populations. Closed responses were analysed using proportions, open responses through identifying repeating ideas and underlying concepts.

Results: Of fifty-two respondents from 25 counties, 87% worked at an investigational site and 75% reported about an interventional study. Studies employed a range of methods to elicit, assess and record participant-reported AEs and related data. Questioning about AEs in 31% of interventional studies was a combination of general (open questions about health) and structured (reference to specific health-related items), 26% used structured only and 18% general only. No observational studies used general questioning alone. A minority incorporated pictorial tools. Rationales for the questioning approach included: standardization of assessment or data capture, specificity or comprehensiveness of data sought, avoidance of suggestion, feasibility, and understanding participants' perceptions. Most respondents considered the approach they reported was optimal, though several reconsidered this. Four AE grading, and three causality assessment approaches were reported. Combining general and structured questions about non-study drug use were considered useful for revealing and identifying specific medicines, while pictures could enhance reports, particularly in areas of low literacy.

Conclusions: It is critical to evaluate the safety of anti-malarial drugs being deployed in large, diverse populations. Many studies would be suitable for contributing to a larger body of evidence for answering questions on harm. However this survey showed that various methods are used to obtain relevant data, which could influence study results. As the best practices for obtaining such data are unclear, anti-malarial clinical researchers should work towards consensus about the selection and/or design of optimal methods.

4.1 Background

When investigating drug effects, study results are influenced by the methods used to collect, assess, record and report outcomes. Studies using different methods can be difficult to evaluate within a body of evidence, so it is beneficial to harmonize conduct (94). There has been significant harmonization of methods relating to efficacy outcomes. For malaria, this entails a World Health Organization (WHO) protocol for therapeutic efficacy evaluations and a categorization of treatment response developed by the WorldWide Antimalarial Resistance Network (WWARN) (96, 198). In general there has been less attention to developing harmonized methods for evaluating safety (monitoring the presence or absence of harm), reflecting its historical marginalization within drug development (64, 65, 199). Notable exceptions include case definitions for adverse vaccine reactions, standard assessment and reporting of adverse effects in rheumatology, HIV and oncology, and instruments for determining specific anticipated effects, for example, extrapyramidal motor side effects of antipsychotics (104, 116, 122, 200, 201). Little work concerns the systematic detection of unanticipated effects (202). For malaria trials, WHO recommends that, to ascertain the incidence of adverse events (AEs), participants be asked about symptoms that have emerged since the previous follow-up visit by "direct questioning" (96).

The manner of ascertaining AEs from subjective participant reports may be more problematic for harmonization compared to those determined through tests or examinations, as it is proposed the former are shaped by memory, expectations, consideration of information required and willingness to report (81, 173, 183). The interpretation and comparison of AE data is further complicated if studies use different methods for assessing severity and relationship to study drug. Research shows that questioning (elicitation) methods can play a role; when asked about specific conditions or body systems as opposed to general questions about health, participants typically report more (139). This suggests certain techniques overcome particular barriers to reporting (183). Some, therefore, recommend using questionnaires or checklists of potential AEs due to their greater sensitivity (131). However, opinion is divided; despite evidence that participants do not always report an AE when asked a general question, some counsel against more detailed questioning so as to prevent inducing a particular response (4, 122). Others suggest that AEs detected by detailed questioning are not as clinically meaningful as those mentioned spontaneously and that general questioning provides a better evaluation

of drug-placebo difference (141, 165). That this initial stage of collecting AE data has been largely excluded from efforts to harmonize safety evaluation methods within some therapeutic areas perhaps reflects these complexities and debates.

Previous medical histories, non-study drugs (previous or concomitant medication), and study drug adherence data, all largely ascertained from participant reports, are integral to assessing harm. Yet these are rarely included in debate about the challenges of obtaining adverse effect data, despite evidence that participants fail to report some medications when asked (25). As for AEs, the questioning tool has been found to influence reports (173, 174).

It has been acknowledged that the above factors inherent in evaluating anti-malarial drug safety and tolerability are challenging, and that there is a need for further guidance (92). To contribute to debate about these concerns a survey was conducted about the methods researchers use to obtain data for evaluating participant-reported harms. A survey was chosen over a literature review as the detail of how data are collected and managed is insufficient in most publications (5). A survey could also prepare for subsequent collaborative work within the anti-malarial research community, with the aim of selecting or designing suitable harmonized methods. This survey was conducted within the ACT Consortium, a group of researchers conducting projects relating to the wide-scale implementation of artemisinin-based combination therapy (37). Several projects contribute safety outcome data to a register coordinated by the Liverpool School of Tropical Medicine (in collaboration with the Malaria in Pregnancy, MiP, Consortium), which can produce individual study reports and pooled analyses, and can also accept data from outside of the ACT Consortium (203). This will be a valuable resource for understanding more about the harmful effects of anti-malarials, particularly as drugs are being widely distributed within high-risk populations who could be quite different to those studied within registration clinical trials.

4.2 Methods

4.2.1 Survey objective

The objective was to explore the methods used to detect, assess and record participant-reported medical history, AE, study drug adherence and non-study drug data in anti-malarial clinical drug research.

4.2.2 Survey population and sampling

Those eligible were anyone involved in the elicitation (by questioning) and recording of harms and related data from participants in any malaria clinical drug study, whether directly interacting with participants, or designing questioning, assessment or data recording methods, or taking responsibility for these tasks. There was no sample size calculation due to the nature of this descriptive survey. Instead, the survey was as inclusive of respondents as possible within the available budget and timelines. Recruitment was active and passive. Personal e-mail invites were sent to known contacts and potential eligible respondents identified from PubMed, clinicaltrials.gov, the Pan-African Clinical Trial Registry and the Initiative to Strengthen Health Research Capacity in Africa's database of African Health Researchers (204-207). The ACT Consortium, MiP Consortium, WWARN, and Global Health Trials publicized the survey through newsletters, webpages or mailing lists, and it was advertised at several conferences (37, 203, 208, 209). During recruitment it was determined that sponsors may not always know the detail required, though would be valuable in subsequent debates, so they were not actively targeted further. Survey respondents were asked to coordinate responses within project teams unless members had differing experiences. Self-selection ultimately determined who participated.

4.2.3 Survey conduct and analysis

The survey (Supplementary File 4.1) was developed in SurveyGizmo® for completion online from August 2012 to January 2013 (210). Questions captured the detail, rationale and application of methods used within various malaria drug study designs and populations. Respondents were also asked about important and feasible approaches over and above what they had experienced. Examples of tools (e g, case record forms) and recommended literature were requested. It was too complex to ask about trial design, such as length of participant follow-up, although these are important aspects of assessing AEs that can also

hinder meta-analyses (5). The questionnaire was piloted with seven eligible respondents, establishing whether questions were understood as intended, and requesting suggestions for improving content and conduct. That the survey was considered acceptable and relevant contributed to face and content validity. The survey was anticipated to take 20 minutes to complete. After recruitment, content was downloaded into Microsoft Excel®. Closed question responses are presented using proportions, and open responses, through mapping underlying concepts and repeating ideas identified (211).

4.2.4 Ethical considerations

The University of Cape Town's Faculty of Health Sciences Research Ethics Committee gave written approval for the survey, and consent to take part was integral with its completion. Access to responses was restricted to those in the investigational team. Once analysis was complete the survey was de-activated, removing links between e-mail addresses and the website.

4.3 Results

4.3.1 Survey respondents and nature of studies described

There were 150 non-duplicate visits to the survey and 56 questionnaires sufficiently completed for inclusion. Four were excluded as they concerned vaccine trials which were not the focus of the survey. Of the remaining 52, there was representation from 25 countries (Figure 4.1). Eighty-one per cent of respondents had more than five years' involvement in malaria clinical research. Most (85%) worked at a study site; in addition there were four representing sponsors/sponsor-investigators, four taking other coordinating roles and an advisory board member. Twenty-five (57%) site respondents were Principal Investigators (PIs), the remainder being co-investigators/researchers (n = 10), study coordinators (n = 7), or other/unknown (n = 2). Forty-three (84%) took responsibility for selecting or developing the methods used to collect participant-reported AE and/or non-study drug data. The majority of clinical research conducted by respondents was non-commercial (77%).

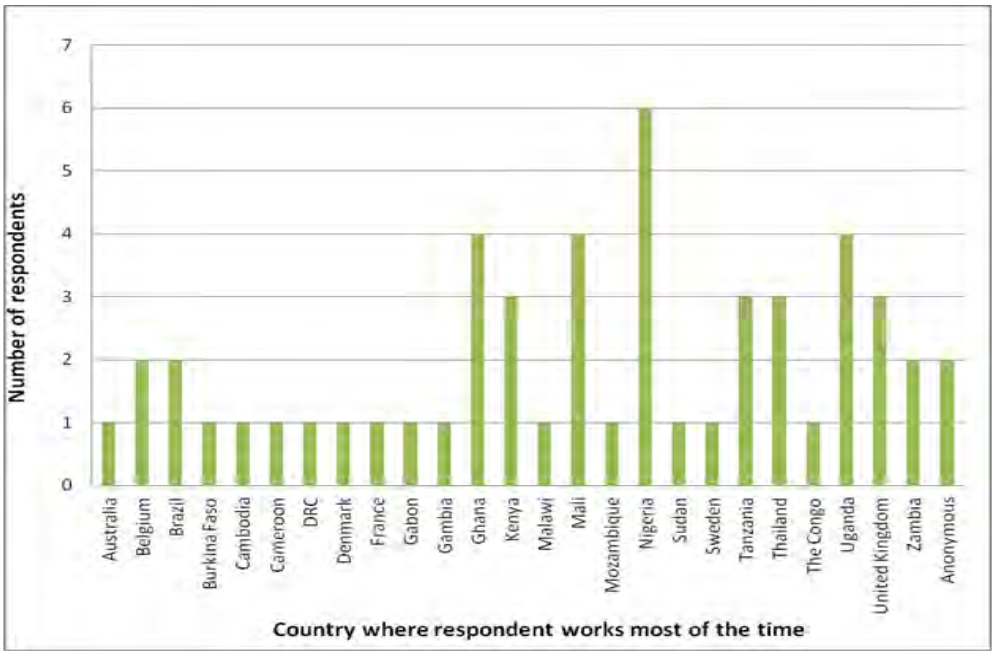

Figure 4.1 Survey respondents by geographical region where respondent works most of the time (N = 52)

Respondents answered the survey with reference to their most recent study, of which 39 (75%) were interventional, nine (17%) observational, the remainder undefined. Studies employed a variety of staff to take a medical history from participants or care-givers, including any change in health or use of treatments (Figure 4.2). Sixty per cent of studies included children between one and 17 years old; the median age in studies where children were asked directly about their health was five years, while for medication-use it was seven years. Forty per cent (20/50) involved a translator in participant conversations. The limited number (n = 5) and diversity of the AE data collection tools submitted meant that they could not be explored usefully. Hereafter N = 52 unless otherwise indicated (not all questions were relevant for, or answered by, all respondents).

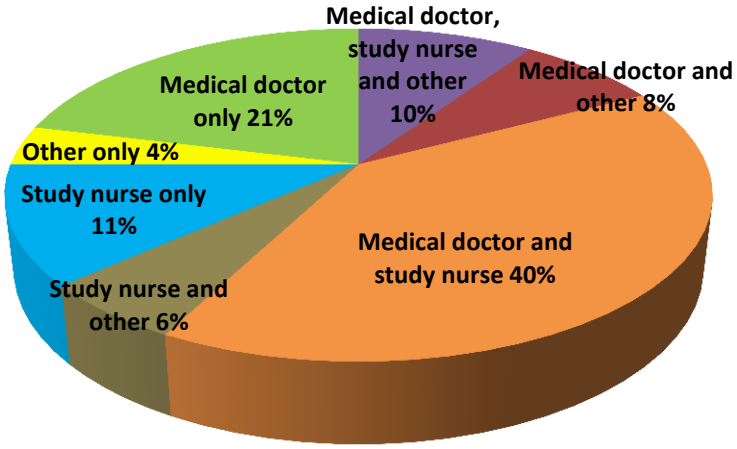

Figure 4.2 Proportion of studies by staff member(s) who question participants about health and drug use (N = 52). "Other" most frequently referred to community health workers, field workers and various trained members of the research team.

4.3.2 Questioning study participants about health to collect AE data

A range of methods were applied to elicit participant-reported AEs. Questioning in 12/39 (31%) of interventional studies was a combination of general (without reference to particular conditions or body system) and structured (with reference to particular conditions or body system), while 10/39 (26%) were structured only and 7/39 (18%) general only (Figure 4.3). All of the nine observational studies described incorporated structured questioning. General enquiries involved phraseology about one of the following concepts (in descending order of frequency by two or more respondents): general enquiry about feeling (e g, "How have you [has your child] been feeling?"), explicit enquiry about change in health (e g, "Have you observed any change or new complaint"), and enquiry with implied causality reference (e g, "Did your child experience any serious side effect from the drug"). One respondent used a phrase possibly aimed at overcoming some impediments to reporting: "Can you confide in us to give a sincere answer". When general phrases were mandatory for use in a study, all respondents were confident the phrase was used as prescribed (N = 12).

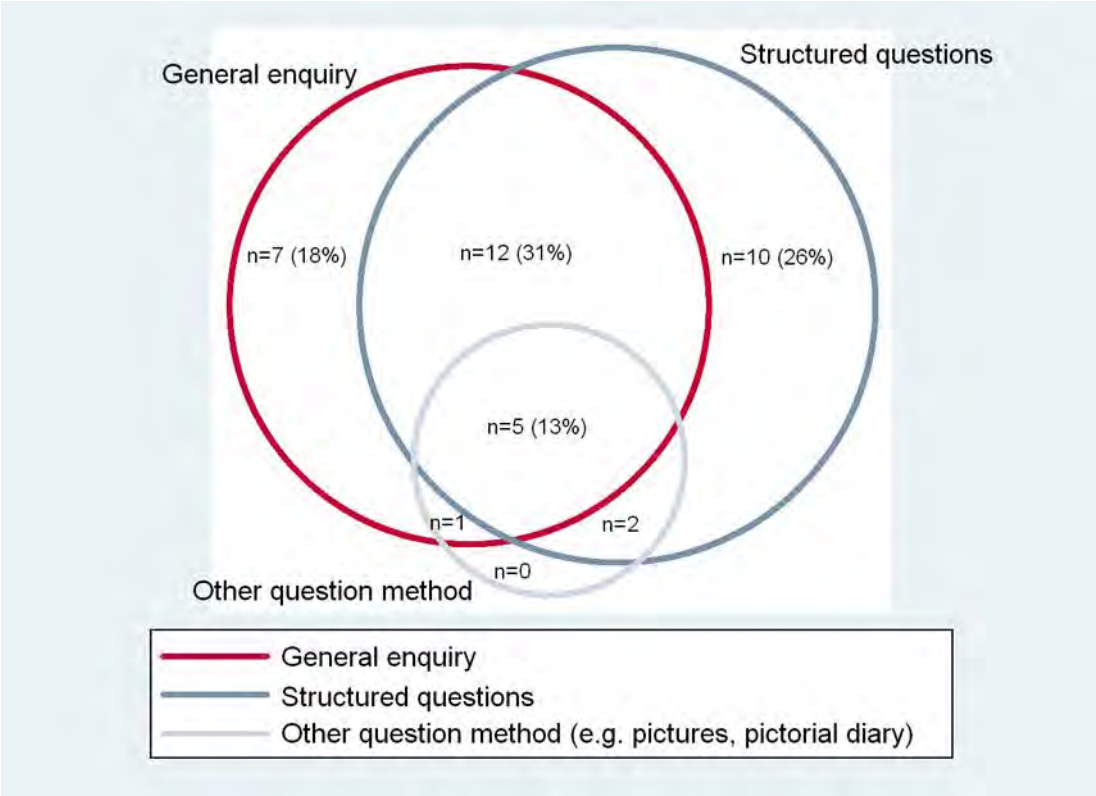

Figure 4.3 Questioning methods about health used to capture AE data in interventional studies (N = 39, 2 = not known).

Structured enquiries involved numerous permutations of symptoms (malaria-specific or not), body systems, and expected adverse drug effects. When used, they were mostly in reference to a prepared list (85%), either exactly as prescribed, paraphrased or a combination. Other than general and structured questioning, eight studies (N = 52, 14%), 5 of which were interventional (N = 39, 13%) incorporated other methods, including pictures and/or pictorial diaries (e g, clinical presentations). Only one such tool was referenced [35].

Key attributes valued in the rationale for method choice were: standardization of AE assessment or data capture, specificity or comprehensiveness of data sought, avoidance of suggestion of a certain response, feasibility and understanding participants' perceptions about health (Table 4.1). Other than more structured enquiries being used to obtain comprehensive reports, there was overlap in the use of different questioning methods to fulfil the same rationales. Most respondents (35/47, 74%) said the approach used in their last anti-malarial study to elicit participant-reported AEs was the one they also considered optimal (important and feasible). Two of these reported that the original method worked

or was reliable, though no details were given on validity or reliability assessments. Five respondents said, however, that the method depended on the needs and circumstances of the study or participant population, and seven suggested an alternative to their originally reported method; one moving from a combination of general and structured questioning to only general questioning (to reduce bias in reporting) and six by proposing a more thorough strategy, incorporating further types of questioning or tools. For the latter, pictures, wall charts, and translations of instruments into local languages were deemed potentially useful for obtaining comprehensive data, and beneficial in areas of low literacy. Seventy-seven per cent (36/50) used the same method of assessing health at baseline as for follow-up visits to detect AEs. Reasons for using a different approach at baseline were to target malaria symptoms and eligibility criteria, including drug-related allergies.

Rationale	Example quotes from respondents	Question method used in the study			
		General enquiry only	**Structured enquiry only**	**General and structured enquiry combination**	**Added pictures, diary, charts, collecting packets, showing sample drugs**
Standardisation of assessments (including historical use of a method in the research group)	*"A systematic approach based on pharmacovigilance procedures developed by our collaborators"* *"We are used to that"*	x	x	x	
Specificity of data (seeking information about particular adverse events, malaria symptoms or drugs)	*"We wanted to find out about specific symptoms and adverse effects"* *"The named drug questions targeted drugs of special interest"*	x	x	x	
Comprehensiveness of data (participant guidance, report clarification, overcoming barriers to reporting such as poor recall or ability to name medicines)	*"To provide a clear understanding about what investigators are looking for and to be sure they capture all complaints from study participants"* *"To get more information which may have been missed during the initial interview"*		x	x	x
Avoidance of suggestion	*"Keeping questions open and not leading so that only events significant to the patient are reported"*	x		x	
Feasibility	*"A simple screen [as the] main focus of the study was not safety/tolerability"* *"Appears simple and not complex"*	x	x	x	
Understand participants' perceptions about health	*"[To] know if [symptoms] are related to chronic disease or traditional belief"*		x	x	

Table 4.1 Rationale for choice of questioning method about health and non-study drug treatment (28/52 respondents)

4.3.3 Assessing and recording AE data

A greater proportion of studies in which AEs were not assessed for severity were observational (63 versus 11% of interventional study reports). Of 20 respondents describing an AE severity assessment method (19 interventional and one observational), 30% used a tool published by the WHO, 20% by the Division of Microbiology and Infectious Diseases, and 10% by the US National Institute for Cancer (116, 200, 212). One respondent commented that the WHO tool does not work well for neutropenia and other variables in their setting, citing Saathoff (213). The remaining 40% (all interventional), used the categories "mild, moderate, severe" (or synonyms), with various definitions largely based on the concept of the AEs' impact on daily living. Methods used to determine the relationship between AEs and the study drug(s), included the WHO-Uppsala Monitoring Centre causality tool, Bradford Hill criteria, and a range of in-house categories or those adopted from unknown sources (214, 215). Forty per cent (19/47) said study participant AE reports were entered verbatim into their database, either solely or with the staff member's or standard terminology for the event. Standard terminologies were largely protocol-specific, though three interventional study respondents cited MedDRA® (216).

4.3.4 Non-study drug use

There were no obvious differences between observational and interventional studies in how non-study drug data were elicited. The majority described general questions (41/47, 87%), English terms being 'medicine', 'medication' or (less often) 'drug'. There was no common practice for whether these phrases were required to be used by study staff or not. Categories explicitly referred to in general enquiries included: prescription-only medicines (32/39, 82%), over-the-counter medicines (27/39, 69%), traditional treatments (24/39, 62%), supplements (11/38, 28%), and vaccinations (8/39, 21%). Sixty per cent (28/47) of respondents used a structured medication enquiry, largely in combination with the general enquiry. About half of these incorporated a study-specific tool specifying treatment class and/or name. There were two reports of using pictures of either a mosquito or drug packets to elicit non-study anti-malarial drug use.

Structured enquires were deemed useful for revealing specific medicines, particularly anti-malarials, and to help name medicines mentioned in response to a preliminary general enquiry. Pictures or diaries were considered to enhance reports and were useful in areas of

low literacy (Table 4.1). Sixty-six per cent (27/41) of respondents were comfortable the approach they used was optimal. Those indicating it could be improved, suggested augmenting a general enquiry with specific mention (or pictures) of medicines such as anti-malarials or drugs potentially interacting with the study drug. One person recommended the content of such lists be developed from pre-study qualitative work in the target population, though another described pictorial tools as "good but logistically hard and…costly to develop". A further proposal was for all staff in contact with a participant to enquire about non-study drug use. There was mention of inherent limitations to questioning about non-study drug intake, "To get medical attention [study participants] lie they have not taken anything", and one respondent suggested assaying blood drug concentrations.

4.3.5 Study drug adherence data

Eighty per cent of reported studies (33/41) collected study drug adherence data using combinations of dosing being directly observed wholly or partially (n = 24), participant recall (n = 14), pill counts (n = 7), dispensing confirmation (n = 8), and pill diaries (n = 2). In 65% or more of responses (N = 32), data were captured regarding quantity of doses dispensed, doses observed, duration of total therapy, time of doses and whether the participant vomited. About half of respondents reported capturing the reason for non-adherence and whether the study drug was taken with food (where applicable). When patients vomited the usual practice was to repeat the dose within a specified period and make a note in the records. Definitions for adherence included taking all doses appropriately, the proportion of participants who took at least a specified percentage of study drug (80 or 90%), and use of the categories adherent/probably non-adherent/definitely non adherent. Seventy-nine per cent said that adherence was described in the study report.

4.4 Discussion

Exploring processes for evaluating harm in anti-malarial drug research may offer insight into how study participants' experiences become facts in databases. This was the first investigation into the detail of methods used within malaria clinical drug research to elicit, assess and record participant-reported harms and associated data. These are critical components of anti-malarial drug safety evaluations, including systematic reviews and

meta-analyses, yet are not readily available in publications (217). Fortunately, this situation is generally improving (218). An online survey was pragmatic and had participation from 25 countries, representing over 50 studies.

Elicitation of AEs

The survey showed that various permutations of general and structured questioning are used to elicit participant-reported anti-malarial AEs, with a minority of studies incorporating pictorial tools. The concern is that differences in questioning methods, even nuances between general enquiries, influence data elicited (4, 139). The few case record forms submitted validated the survey findings that specific general enquiry phrases are obligatory in some studies. While conversations are unlikely to follow a script completely, a preliminary standard phrase can orientate staff and participants to the intention of questioning. Structured questioning for AEs with reference to predefined fields was frequently used to permit standardized assessments and data capture, and obtain comprehensive replies. While structured enquiries can increase the number of AEs reported (hence burden on the study team) they can also facilitate a tolerability assessment, which is particularly important at a population level (92). The survey results reflect the literature that children from five to seven years old can contribute valuably in medical consultations. As young children may be marginalized in conversations, however, it is proposed that study staff are guided in managing triadic communications (219).

Most respondents considered the questioning approach they used was optimal. However, the subtle shifts observed between rationale for methods reported in use and respondents' perceptions of optimal approaches indicate reflection on appropriate methods, and particularly the potential value of novel methods. In other therapeutic areas, work is underway to harness techniques used in developing validated patient-reported outcome measures (PROs) for AEs, a trend being self-report via the internet (83, 220). Many existing instruments, however, restrict the development of instrument items (i. e. potential AEs) to those side-effects already associated with a drug by patients and/or experts, and generate summed scores rather than individual AE reports (129). This may not allow for adequate detection of unexpected AEs or facilitate a traditional safety evaluation through incidence measures. Nevertheless, finding ways to better represent a participant's experience of an AE is important. For malaria it makes sense to pursue methods suitable for areas of lower

literacy, including pictures, mobile phones, or innovative methods being explored in other areas of health care communication, such as when patients indicate health problems on body outlines kept at home for later discussion at the clinic (74, 221).

The survey results did not substantiate whether variation in AE enquiry methodology is a factor of study design. This issue should be debated; for instance, whether a more sensitive enquiry is used to detect AEs in earlier phase trials, compared to later phase studies where a drug has a better understood safety profile. The caveat, however, is uncertainty about which questioning method(s) produce valid data and whether there should be consistency throughout study visits; while it is important to avoid leading patients to a particular response, it may also be inappropriate to use a less sensitive measure to detect change in health post-intervention than one used to determine a baseline. There should be further methodological research to investigate these concerns, including qualitative work in both participants and staff about understanding of terminology and meaningful outcomes in malaria-endemic populations. The ACT Consortium plans to contribute through a Cochrane review of research comparing adverse effect elicitation methods (222). Available data could potentially be distilled to relevant concepts for local adaptation according to the study design and population.

Grading severity of AEs

A variety of AE severity/toxicity grading tools and methods for determining relationship with study drug were reported. Researchers in other therapeutic areas regularly debate the development of disease-specific toxicity criteria for end points important to health practitioners and patients (122). Should malaria researchers do the same, such criteria should be based on locally relevant reference ranges for laboratory parameters (223). Staedke et al. observed variability between sites in assessing AEs for severity and causality when using standard grading scales, reflecting differing value judgements of staff and participants (92). There is a need then to think carefully about terminology and application of criteria so that there is consistent interpretation (224). Similarly, because of the plethora of causality assessment criteria used in clinical research, there have been recommendations of simpler categories (e g, related/not-related; reasonable possibility/no possibility of attribution), with more emphasis on determining causality by statistical methods where possible (4, 224).

Eliciting and recording study and non-study drug data

Survey respondents considered combining structured and general questions about non-study drug use useful for revealing and identifying specific medicines. Pictures, diaries and samples etc., were seen to help with recall and participant understanding of what to report, particularly pertinent for areas of low literacy. The cost of formative work for development of such tools may be prohibitive for individual study teams, but could be shared with other groups recruiting from a similar patient population (even if studying different therapeutic areas). Often daily anti-malarial doses can be fully observed. However, when required (e g, for twice daily dosing of artemether-lumefantrine), there were various methods reported as being used for measuring adherence, whether indirect and objective (pill counts and dispensing information) or indirect and subjective (diaries and recall). Combining recall questionnaires or pill diaries with an electronic pillbox and count may provide the most accurate information, though in areas of low literacy or limited study budget a recall questionnaire with pill count may be more realistic (WWARN, personal communication, 2012). More work is needed to define adherence for the variety of drugs or combinations used in malaria to inform meta- or pooled data analyses.

4.5 Limitations

Surveys inherently have constraints in terms of their reach, respondents' understanding of questions and their ability or willingness to give comprehensive answers. It is assumed that responses reflect the methods used in practice. However, it would have strengthened the study if other field staff were included to know whether their responses differed from those of the PIs, co-investigators, study coordinators and researchers who took part. The survey results do not represent all anti-malarial drug clinical research studies and it was not possible to detect differences in methodologies between study designs or populations. It is hoped that any such limitations will be overcome in future debates. In particular, more representation from sponsor organizations and regulatory authorities will be key.

4.6 Conclusion

Many anti-malarial drug studies would be suitable for contributing to a larger body of evidence for answering questions on harm, regardless of design or whether safety was the

primary objective (225, 226). Meta-analysing and pooling data can increase the power to generate new signals, and clarify the incidence and possibly drivers of known harms. Applying these tools is increasingly urgent as large populations, some who may have multiple co-morbidities or be taking concomitant medications, are exposed to anti-malarial drugs. Should anti-malarials be distributed in asymptomatic or malaria negative populations it is particularly important to evaluate benefits versus harm (227). Studies evaluating strategies such as intermittent presumptive treatment, screening and treatment, and mass drug administrations should assess safety, including the less severe adverse effects as tolerability can impact on community acceptance of, and individual adherence to medicines (15). In order to synthesize study results effectively researchers need to move towards harmonized methods for obtaining safety data. As there are many unanswered questions about the best practices for eliciting, assessing and recording such data, priority areas for further work should be established through dialogue within the anti-malarial clinical research community. This could include 1) nesting methodological research within studies to find optimal approaches or tools for questioning participants to obtain AEs and related data, 2) deciding whether to adopt an existing toxicity grading scheme from another therapeutic area or develop one for malaria endemic populations, or 3) developing guidance on use of a common causality assessment tool. User-friendly open access databases suitable for a range of study designs could also be developed collaboratively to help researchers manage their data efficiently (228). Where appropriate these should incorporate harmonized fields and terminologies so that they may be used more widely in the non-commercial sector. The ACT Consortium plans to contribute by facilitating a Delphi process; a summary of the available literature will be presented to interested malaria clinical researchers who will then work towards consensus about the appropriate design of relevant and feasible methods to detect these important data. This work could also catalyze much needed progress in the detection and investigation of harms in other therapeutic areas, improving the ability to compare or synthesize studies (229).

CHAPTER 5: HOW EXPERIENCES BECOME DATA; THE PROCESS OF ELICITING ADVERSE EVENT, MEDICAL HISTORY AND CONCOMITANT MEDICATION REPORTS IN ANTIMALARIAL AND ANTIRETROVIRAL INTERACTION TRIALS

Elizabeth Allen[1]; Adiel Mushi[2]; Isolide Massawe[3]; Lasse Vestergaard[4]; Martha Lemnge[3]; Sarah Staedke[5], Ushma Mehta[1], Karen Barnes[1], Clare Chandler[6]

[1] Division of Clinical Pharmacology, Department of Medicine, University of Cape Town, South Africa.

[2] National Institute for Medical Research, Dar es Salaam, Tanzania.

[3] National Institute for Medical Research, Tanga Centre, Tanzania

[4] Centre for Medical Parasitology at Institute of International Health, Immunology and Microbiology, University of Copenhagen, and Department of Infectious Diseases, Copenhagen University Hospital, Copenhagen, Denmark

[5] Department of Clinical Research, London School of Hygiene & Tropical Medicine, U.K

[6] Department of Global Health & Development, London School of Hygiene & Tropical Medicine, U.K

Publication status:

Published: BMC Medical Research Methodology. 2013 13(1):140.

Author contributions

The candidate (EA) conceived of the study, led design, data collection and interpretation, conducted South African interviews and wrote the paper. CC had significant input throughout all these stages. AM and IM conducted Tanzanian interviews. AM refined study tools and supervised the qualitative team in Tanzania. KB, SS and UM contributed to study design and data interpretation. ML and LV supervised the Tanzanian team, LV also contributed to data interpretation. All authors contributed to and approved the final paper.

Abstract

Background: Accurately characterizing a drug's safety profile is essential. Trial harm and tolerability assessments rely, in part, on participants' reports of medical histories, adverse events (AEs), and concomitant medications. Optimal methods for questioning participants are unclear, but different methods giving different results can undermine meta-analyses. This study compared methods for eliciting such data and explored reasons for dissimilar participant responses.

Methods: Participants from open-label antimalarial and antiretroviral interaction trials in two distinct sites (South Africa, n = 18 [all HIV positive]; Tanzania, n = 80 [86% HIV positive]) were asked about ill health and treatment use by sequential use of (1) general enquiries without reference to particular conditions, body systems or treatments, (2) checklists of potential health issues and treatments, (3) in-depth interviews. Participants' experiences of illness and treatment and their reporting behaviour were explored qualitatively, as were trial clinicians' experiences with obtaining participant reports. Outcomes were the number and nature of data by questioning method, themes from qualitative analyses and a theoretical interpretation of participants' experiences.

Results: There was an overall cumulative increase in the number of reports from general enquiry through checklists to in-depth interview; in South Africa, an additional 12 medical histories, 21 AEs and 27 medications; in Tanzania an additional 260 medical histories, 1 AE and 11 medications. Checklists and interviews facilitated recognition of health issues and treatments, and consideration of what to report. Information was sometimes not reported because participants forgot, it was considered irrelevant or insignificant, or they feared reporting. Some medicine names were not known and answers to questions were considered inferior to blood tests for detecting ill health. South African inpatient volunteers exhibited a "trial citizenship", working to achieve researchers' goals, while Tanzanian outpatients sometimes deferred responsibility for identifying items to report to trial clinicians.

Conclusions: Questioning methods and trial contexts influence the detection of adverse events, medical histories and concomitant medications. There should be further methodological work to investigate these influences and find appropriate questioning methods.

5.1 Background

Assessment of drug harm and tolerability rely, in part, on clinical trial participant reports of adverse events (AEs), medical histories and concomitant medications. However, there is no consensus regarding the detail of how such reports should be elicited, in particular how participants should be questioned about ill health and concomitant treatment use. Heterogeneity in elicitation methods provides potential for measurement error if questioning methods are sub-optimal, and undermines meta-analyses of adverse effects (60, 167). Staff may use general enquiries to identify AEs, such as 'How have you been feeling?', yet the impact of variation in phraseology between or within trials is unclear (4). Some trials elicit responses using detailed, structured approaches like symptom lists, often aimed at ascertaining anticipated adverse effects or AEs of special interest, though detection of unanticipated effects may remain unsystematic (135, 182). Questioning may also involve self-completed forms or diaries. There is evidence that more detailed elicitation techniques increase the sensitivity of participant-reported AEs (139). However, their effect on the nature of reports is unclear. Barber and Santanello (141) found that only what patients considered more bothersome was reported spontaneously when compared to use of a checklist of possible events. This suggests that spontaneously reported AEs may be more clinically meaningful, and supports concerns that detailed methods produce 'noise': clinically irrelevant AEs that cannot be distinguished from background rates (139), (112) Wernicke et al. (165) have also proposed that spontaneous reporting provided larger drug-placebo differences more often than solicitation. There is a dearth of research about methods for the elicitation of previous or concomitant medications in trials, despite evidence that participants fail to report use of other antimalarials when asked (25).

A major challenge in determining the best way to elicit these data is the lack of a gold standard to assess validity of responses; an absolute measure of what patients experience is likely unachievable. However, in other areas of pharmacoepidemiology, including case–control or cohort studies and administrative databases, there has been methodological investigation regarding the accuracy of self-reported past medical conditions and treatments, through comparison with medical or prescription records (173). In those contexts, recall of medical history appears dependent on the type of condition, its perceived importance and a willingness to share information. Pattern of use is influential in recall of past medications, and indication- or medication-specific questions increase prevalence estimates compared to open-ended questions (174).

Factors that shape reporting of beneficial or adverse effects of pharmacotherapy is a poorly researched area (57), although a variety of relevant factors have been proposed in different research areas. These include cultural variations in how health issues are perceived (24); negative emotions (230); beliefs about the medication (231); response-shift, whereby ill health is not viewed as such anymore (79); and gender or language differences that impede communication between patients and health workers (232-234). Belief that drugs must cause certain [side] effects as part of the therapeutic process may also play a role (58). Specifically within the trial context it has been proposed that there may be fear of lack of confidentiality, stigma or negative repercussions as a result of reporting AEs (4, 235). Nocebo effects (harmful response to an inactive product) may be related to information about anticipated effects given to participants during the consent process (82, 236).

This study aimed to identify factors shaping reports of medical histories, AEs and treatments by participants who had malaria and/or HIV and were enrolled in trials of concomitant antiretroviral (ARV) and antimalarial treatments. This is to inform elicitation practices for improving reporting in these clinical contexts. The ACT Consortium (37) is a global research partnership aiming to answer key questions on the delivery of artemisinin-based combination therapy (ACT) for the treatment of malaria. The Consortium involves a number of projects that elicit safety data from trial participants, providing an opportunity to investigate factors influencing data in areas where, due to high rates of illiteracy, it would be challenging to use self-completed questionnaires. This methodological study was nested in two open-label antimalarial and antiretroviral interaction trials (referred to as parent trials hereafter) that both aimed to assess the safety and blood drug concentrations (pharmacokinetics) of the widely-used antimalarial drug, artemether-lumefantrine (AL). The two different study designs and contexts provided an opportunity to explore the influence of a variety of factors on reporting. A small South African trial investigated the use of AL in HIV-infected volunteers, while a larger Tanzanian trial investigated AL in patients infected with malaria and/or HIV.

5.2 Methods

5.2.1 Study design and objectives

This was a mixed method study. Parent trial participants were asked about ill health and treatment use (in order to record AEs, concomitant medicines and medical history) through

a variety of elicitation methods: general enquiries, checklists, in-depth interviews and focus group discussions. The primary objectives were to explore qualitatively participants' experiences of illness and treatment, their reporting behaviours, and their responses to different questioning methods. Trial clinicians' experiences with eliciting and recording participant-reported data were also explored for their perspective on these interactions. Secondly, the study aimed to compare the number and nature of data obtained through use of different question methods. It was assumed that the parent trials should seek medical history data sufficient to determine eligibility and inform the causality assessment of AEs. In addition, they should detect changes in health or treatment use from baseline such that within-trial or pooled statistical analyses could establish signals for, or evidence of harm or poor tolerability, or robust evidence for the absence of harm within the study population.

5.2.2 Study population

The study population was participants and trial clinicians involved with the two ACT Consortium parent trials. In South Africa, this was 18 trial participants between November 2009 and February 2010, in Tanzania, 80 of 832 trial participants between October 2010 and May 2011.

Both trials investigated responses to AL, although with different designs and participant populations. The South African site was an urban tertiary hospital serving a socio-economically diverse population. In this pharmacokinetic and safety trial, groups of otherwise healthy HIV positive volunteers were recruited and admitted for four days to a pharmacology research ward. Prior to this, participants had been followed-up after a single dose of AL. At the time of this nested study they were returning for a multiple-dosing period of AL twice daily for three days, which involved intensive pharmacokinetic sampling (to determine drug levels of AL), plus follow-up on an outpatient basis until Day 21. The Tanzanian site was a rural district hospital serving a predominantly very poor population. In this efficacy, safety and pharmacokinetic trial, HIV positive or negative patients presenting with malaria symptoms were recruited from routine clinics and attended trial consultations individually as out-patients until Day 42. Those positive for malaria received the same six-dose AL regimen as for South Africa, while those testing negative were managed according to their alternative diagnosis.

In both sites, participants who reported additional information in response to checklist questioning compared with a general enquiry were eligible to participate in in-depth interviews and subsequent focus group discussion. This was to explore their disparate responses to different questioning methods. Trial clinicians from both sites were also invited to a focus group discussion once the nested study was complete.

5.2.3 Elicitation methods

Three elicitation methods were compared sequentially: general enquiry, check list and in-depth interview. The general and checklist enquires were conducted on the same day at the baseline and a follow-up visit; the in-depth interviews were conducted within a week of the follow-up visit. See Supplementary Files 5.1 and 5.2 for the question guides used.

General enquiry

General open-ended questions (without reference to particular conditions, body system or medications) about health and treatment use at baseline and follow-up visits were routine in both trials. At the first visit, medical and treatment histories were obtained using open ended questions. However, as the South African participants were returning for a second dosing period, these were recorded in the trial files as AEs and concomitant medications, unless they pre-dated the parent trial. Responses were probed according to common clinical practice in eliciting a medical and treatment history, which is based on the opinion of the enquirer regarding what further details would be required.

Checklists

At the first visit and a follow-up visit (three to seven days post enrolment), the general enquiry was immediately followed with a checklist of questions asked by the trials' clinicians focusing on body systems, symptoms, diseases and treatments. A majority of fields in the checklists were common to both trials (Supplementary Table 5.1) although they could not be harmonised fully. This should not have affected this study, with its primary focus on the experience of differential reporting rather than validity of the question methods. Data reported for the first time in response to checklists were probed as for the general enquiry, and recorded in separate fields so that it was clear which question method detected which data.

In-depth interviews

Of those participants who reported additional information in response to the checklist compared to the general enquiry, a convenience sample was invited to an in-depth interview to explore their disparate responses (described below). These interviews also provided an opportunity to elicit further relevant trial data using a prompted narrative of the participant's trial experience, reflection on previous ill health and treatments, and photographs of typical over-the-counter and traditional medicines available to the study populations. These data could then be compared with data from the other elicitation methods.

5.2.4 Evaluation of reporting behaviours and elicitation methods in context

Qualitative experiences of trial participants were explored through in-depth interviews and focus group discussions.

In-depth interviews

Twenty-seven participants who reported additional information in response to the checklist elicitation method compared to general enquiries were invited to participate in an in-depth interview. The informed consent document explained that the interviewer would talk about participants' experiences of their health and use of treatments but not that the interviews were specifically looking to explore differences in reporting. The interviews lasted one to two hours and were conducted away from the trial venue by a local qualitative researcher who understood concepts relating to collecting trial safety data. After attempting to discover further relevant data using the narrative method described above, participants were asked for help in exploring why their reporting differed by question method, with reference to their own trial data and reports.

Focus group discussions

The same participants who had participated in the in-depth interviews were invited to participate in a focus group discussion within a month of completing the parent trial. The rationale for using focus group discussions was the expectation that further information

about the same concepts may be revealed after reflection in a group context. In addition, unlike the in-depth interviews, information from focus group discussions could not influence their ongoing trial participation, therefore participants may have felt more comfortable talking about particular information that they had chosen not to report during the trial. Two focus group discussions were held in each country. For each (with four to eight participants, separated by HIV status to encourage openness) the facilitator explored treatment use, the meaning or importance of AEs, barriers to reporting and recommendations for improving accuracy and completeness of clinical trial data.

One focus group discussion was also held with the trial clinicians from both parent trials in March 2011. Topics included the appropriate level of data elicitation needed for adequate assessment of harm and tolerability, the relative merits of questioning methods, hypothesised barriers to accurate reporting, and suggestions for improvement. See Supplementary File 5.3 for the question guide.

Question guides for the participant in-depth interviews and focus group discussions were piloted and developed iteratively as data emerged. Tanzanian question guides were translated into Kiswahili using a forward-backward translation from English. EA led South African discussions in English, assisted by English-isiXhosa or Shona speaking social scientists depending on participants' home languages. Tanzanian discussions with trial participants were conducted in Kiswahili by AM and IM. EA and AM conducted the focus group discussion with trial clinicians in English.

5.2.5 Data management and analysis

AEs, medical histories and medications elicited from participants in the sampling frame, including additional reports from those who attended in-depth interview, were described statistically by elicitation method. The nature of reports was considered in relation to trial eligibility for medical and concomitant medication histories, and severity for AEs. Audio recordings were transcribed in the original language (English, isi-Xhosa or Kiswahili), translated into English where necessary, checked for quality and imported into NVivo 8 (QSR International, 2009(237)) with summaries and observations written by interviewers directly after each interview. Transcripts were analysed thematically in terms of

explanations given for differential reporting by question method, and how participants expressed themselves. Relevant text was examined by EA for repeating ideas, which were labeled and grouped into themes reflecting the underlying meaning or concepts behind statements (211, 238). The emerging coding structure was revised after each transcript, with on-going review by CC. EA and CC explored the emerging themes in relation to broader theories.

5.2.6 Ethical considerations

Approval was obtained from the ethics committees of the University of Cape Town, Faculty of Health Sciences and the Tanzanian National Institute of Medical Research Coordinating Committee. Trial participants were told who would have access to their responses and that refusal to participate was without disadvantage for subsequent care or place in the parent trial. Informed consent processes and forms were available in English and local languages. Transport reimbursement and refreshments were provided.

5.3 Results

5.3.1 Study participants

The characteristics of participants who were asked both general and checklist enquiries and the subgroup who took part in the in-depth interviews (and subsequent focus group discussions) are given in Table 5.1. None of the parent trial participants who were approached refused an in-depth interview; however some interviews could not be scheduled due to participant or interviewer availability.

Table 5.1 Characteristics of trial participants

	All participants from the sampling frames		Subgroup of participants interviewed	
	South Africa (n = 18)	**Tanzania (n = 80)**	**South Africa (n = 11)**	**Tanzania (n = 16)**
Number (%) female	13 (72.2)	53 (66.3)	8 (72.7)	9 (56.3)
Median (IQR) age in years	37.1 (33.4 - 39.7)	40 (32.0 – 36.5)	38.1 (35.0 – 42.1)	34.5 (24.5 – 48.5)
Number (%) HIV positive	18 (100.0)	69 (86.3)	11 (100.0)	10 (62.5)
Number taking ARVs	18 (100.0)	63 (78.8)	11 (100.0)	9 (56.3)
Highest education completed number (%)	Data not available for South Africa		Unknown for 1 participant in Tanzania	
None/incomplete primary		26 (32.5)	0	6 (37.5)
Primary school		53 (66.3)	3 (27.3)	10 (62.5)
Secondary school		1 (1.3)	4 (36.4)	0
Higher education		0	3 (27.3)	0

5.3.2 Safety data reports by elicitation method

Of the 18 South African parent trial participants, 16 attended both study visits, of whom 15 (94%) reported differently between the general enquiry and checklists. Of 80 Tanzanian parent trial participants in the trial at the time of the nested study, 76 attended both study visits, of whom 65 (86%) reported differently between the general enquiry and checklists. For practical reasons the final sample size of those who participated in in-depth interviews were 11 in South Africa and 16 in Tanzania. Table 5.2 summarises the reports of adverse events, medical histories and concomitant medications elicited by question method.

Table 5.2 Summary of trial participants' reports elicited by question method

	South Africa			Tanzania		
	Medical histories	**Adverse events**	**Treatments**	**Medical histories**	**Adverse events**	**Treatments**
Number of reports by general enquiry*	4	23	17	285	6	196
Additional number of reports by checklists (% change from general enquiry)*	8 (100.0)	20 (87.0)	23 (135.3)	245 (86.0)	1 (16.7)	2 (91.3)
Additional number of reports by interview **(% change from general enquiry and checklist)	4 (33.3)	1 (2.3)	4 (10.0)	15 (2.8)	0 (0)	9 (4.5)

* All participants in the sampling frame attending both visits (South Africa n = 16, Tanzania n = 76).
**Subset of participants who took part in in-depth interviews (South Africa n = 11, Tanzania n = 16).

There was an overall cumulative increase in number of reports from general enquiry, through checklists, to in-depth interviews with the largest increase in reports found in the checklists when compared with the general enquires. All additional AEs reported through use of the checklists or in-depth interview were rated as mild and unlikely to be related to the trial drug, and no additional medical history or medications detected in using those methods changed eligibility for the trials. While there were large differences in the numbers of reports between the trial sites, it is not informative to make a direct comparison as, apart from the four-fold difference in sample sizes, the type of participant and trial designs were different. It is to be expected that otherwise healthy HIV-positive volunteers in South Africa have significantly fewer medical history reports compared to those in Tanzania who were presenting with multiple malaria symptoms; the South Africans also reported more AEs as they had a second opportunity to do so.

In the more detailed description of data reports (Supplementary Tables 5.2 and 5.3) it was observed that, in particular, baseline symptoms of nausea and vomiting in Tanzania appeared to be under-reported until they were specifically asked about. Similarly, body pain and fatigue were not reported until after specific questioning. However, this may be because they were captured within other terms (e.g. tingling or painful sensation, abdominal pain, chest pain, muscle pain and joint pain for body pain, and weakness for fatigue) in the Tanzanian questioning tool. Night sweats may have been reported within the concept of fever.

5.3.3 Results from the trial participants' narratives

Experiences of their health or use of treatments become data through a process including participants' recognising the information required, considering a reply, and effective articulation, or naming, of the response (Figure 5.1).The underlying social contexts could help to explain how specific barriers and facilitators to participant reporting manifest in these trial sites.

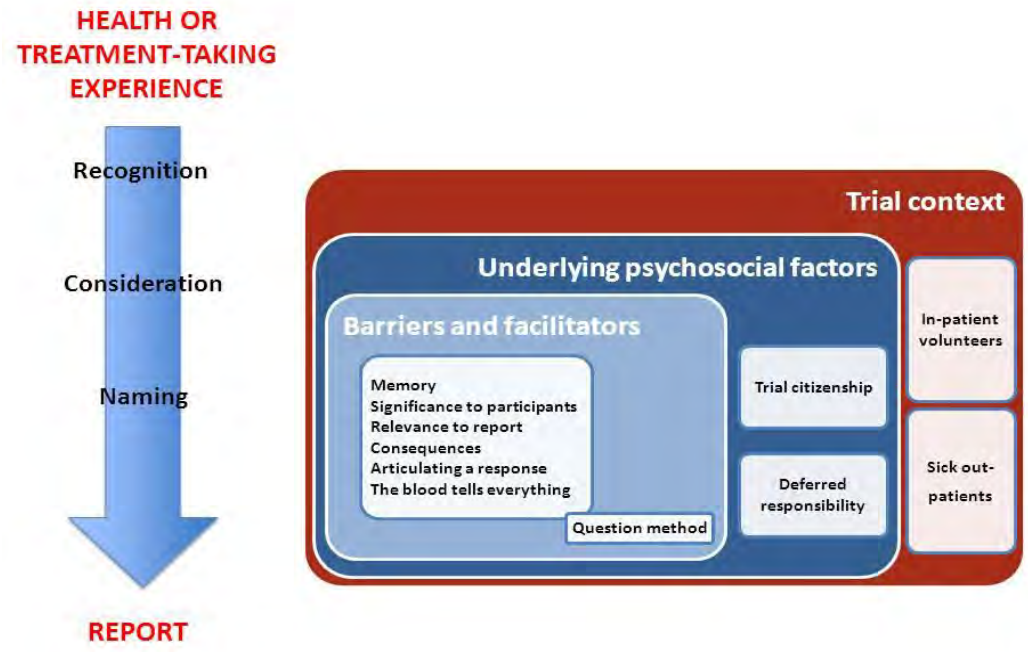

Figure 5.1 Diagram of trial participants' narrative responses

5.3.3.1 Factors shaping participant reporting of health and treatments

Memory

When discussing discrepancies between their responses to the different questioning (elicitation) methods, participants often said they forgot to report things in response to the general enquiry and, to a minimal extent, the checklist. Minor, intermittent and resolved health issues were more likely to be forgotten compared to those more severe, persistent or current. There were no consistent categories of treatments recalled more easily, except

for ARVs, which could never be forgotten as they are considered so important to sustaining health:

> "*I only used it [diclofenac, an anti-inflammatory] to reduce the tooth pain. When the pain stopped I [forgot] them. But with these ones [ARVs] I have no time to [forget] them, maybe God takes my spirit [fear of ill health/death from non-adherence]*" [In-depth interview (IDI) 11, Tanzania]

Participants predominantly attributed forgetting to report to the burden of having too much to remember, although two South Africans said it was due to significant memory decline since starting ARVs. The detail of the checklist, and to some extent the in-depth interview, was strongly viewed as helpful in overcoming memory problems.

Significance to participants

When faced with a general enquiry, participants in both sites described conscious decisions to report current medical histories and AEs that were more bothersome, severe or persistent as opposed to those experienced as intermittent or less severe. Several South African participants also described decisions to first spontaneously report a new deterioration in health during the trial only when the symptoms worsened further. South African participants often recounted medical history in relation to when they initiated ARVs, when health was at a low point from AIDS-related illness or initial side-effects. Illnesses experienced since this period (major or minor) were commonly under-reported,

> *"Ever since 2006 I don't want to lie but I don't feel anything… I always just get sick like ordinary people"* [IDI 4, South Africa]

This extended into reporting about AEs during the trial period, with declarations of good health followed later by statements like *"No, I was not sick at all. Just only coughing, flu, only that"* [IDI 8, South Africa]

Participant definitions or normalising of ill-health influenced when and how they revealed information in the in-depth interview. Categorising illness as 'major' or 'bad' was frequently associated with hospitalisation; tuberculosis treated as an outpatient was not being 'badly

sick', and only revealed in response to a specific question about tuberculosis. One participant did not mention her blindness in one eye until specifically asked.

Non-significance amongst Tanzanian participants sometimes reflected a slightly different concept. Rather than relating to classification of an experience as 'normal' and therefore not reported, participants seemed to shape their perception of significance more around expectations of what would be significant to the doctor. When asked about symptoms by way of a general enquiry, those deemed the main problem(s) were chosen with expectation that others would be treated by default:

> *"I don't think that you can tell the doctor one part after another that is in pain. … But you may decide to tell him the basic problem, or what is making me more sick. If I tell the doctor that I have fever he might give me the medicine … then all that I am feeling will calm down"* [Focus group discussion (FGD) 1 Tanzania respondent 5]

Treatment use phraseology meanwhile revealed a hierarchy; after ARVs or antimalarials, use of intermittent or over the counter substances (such as painkillers or vitamins) were mentioned, but qualified with 'only' or 'apart from'. These perceptions of significance may intersect with the next factor shaping reporting behaviour: relevance to report.

Relevance to report

Participants appeared to delay reporting experiences that they perceived irrelevant, with the checklists helping them to decide what was necessary. Manifestation of relevance to report was, largely, different between sites. In South Africa, decisions to report were sometimes related to the trial's objectives, i.e. whether the information would contribute to the success of the trial:

> *"If I say after drinking the tablet I felt weak and tired … then they will write it … [as] the effects of the tablet. [To] them it gives …the wrong impression"* [IDI 5 South Africa]

History of sexual dysfunction, and poor memory (not on the checklists), were only revealed in the in-depth interview by two participants, who described these as irrelevant to the trial,

and also untreatable, and therefore not useful to report. In Tanzania, relevance was chiefly attributed to the medical context. If a symptom at baseline was considered by the participant to be due to something other than malaria, such as work-related activities, it might not have been reported unless specifically prompted for. Similarly, if medication was bought for something other than malaria there was no reason to report:

> "*I thought because I purchased it [pain killer] by myself because of the tooth pain I had no reason of telling him. But the Fansidar I purchased, I had a reason of telling him, because I can be tested and seen with malaria again. So they'll understand, after the Fansidar being useless, what [they should] do*" [IDI 11 Tanzania]

Consequences – attempts to control a situation

Participants in both sites described themselves or others in the trial withholding information in response to general questions and checklists for fear of negative consequences. These negative consequences were starkly different between sites. In South Africa, participants feared exclusion from the trial; there were three explicit, independent, second-hand reports of self-treatment with laxatives, non-adherence with ARVs, and gastro-intestinal illness that were withheld for this reason. When such scenarios were presented during the focus group discussion, participants denied that this happened, saying that misreporting could undermine the trial's objectives. In contrast, Tanzanians considering consequences of reporting (or not) did not mention the trial; they were worried about going against hospital 'rules'. One said she did not report an anti-inflammatory taken during the trial for this reason. Most discussion on this topic in Tanzania, however, was in the focus group discussions and concerned fear of reporting traditional medicines:

> *"We who are living with [HIV], .. we are highly advised to refrain from traditional medicine issues because there is a difference between traditional medicines and these drugs. ….Yesterday I felt ill and went to dig muarobaini [neem tree root used for malaria] but without feeling any relief after drinking it. If I come to [doctor] today and he asks me what kind of medicine you used yesterday, telling him muarobaini….I must lie because already that is wrong. …That is due to fear of telling the doctor because we were already prohibited that thing. …There is some sort of [difficulty] telling the truth though it's true, but the truth remains within me."* [FGD 1 Tanzania, respondent 3]

Articulating a response

South African participants, who generally did not take long-term medications apart from ARVs, appeared to have a small personal formulary of named over-the-counter medications, but names of ad-hoc prescription medications were largely unknown. Though literate, one man displayed little interest in knowing what his prescription was. By contrast they had impressive knowledge of previous and current ARV regimens. In Tanzania, names of prescription medicines, including ARVs, were not known. As for HIV itself, ARVs were mentioned euphemistically, even when reverence appeared as high. In focus group discussions, Tanzanian participants initially cited their lack of education and an inability to read English labels, though it became clear that names of ARVs and other long-term prescription medicines are seldom verbalised during prescribing and dispensing. Meanwhile, equally complex names of antibiotics and antimalarials are known, because they are talked about in public.

> *"Respondent: But I have not kept those [ARVs] in my mind. I have their container but I have never read it.*
>
> *Interviewer: How come you knew amoxycillin?*
>
> *Respondent: [laughs] They are simple to pronounce*
>
> *Interviewer: How did you get that name?*
>
> *Respondent: Because they are mentioned in the streets … like aspirin, simple names".* [IDI 4 Tanzania]

The blood test tells doctors everything they need to know

When South African participants narrated what happened in the trial, this was largely about blood tests and examinations, with little mention of the verbal discussion with the clinician until they were prompted. While this could relate to the numerous blood samples taken in the trial ward and knowledge about the trial's pharmacokinetic objectives, it persisted in their descriptions of follow-up visits, when tests were far fewer. When asked about the relative importance of blood tests versus being asked about their health and use of medicines, several declared the former as equivalent or superior.

"*[The test is] like they are asking about your health. … it's the same thing*" [FGD 01 South Africa respondent 01]

Others said blood could reveal non-adherence to ARVs and use of any substance, including traditional medicines. Tanzanian participants did not display such a marked pattern, possibly because they expected to be asked about malaria symptoms and what antimalarials they had tried in order to inform treatment. However, when questioned about the importance of blood tests versus being asked about their health and use of medicines their discourse reflected that the former were considered a far superior source of information for the doctor. One participant indicated that questions were asked merely for the doctor to identify the more important test required.

5.3.3.2 Social context of the trials

While there may be influence from psychological factors on the factors shaping reporting, conversations with participants in this study reveal social constructs that may underpin reporting behaviour. A minor construct "Being a subject" was common to both sites, though much stronger in South Africa; participants indicated they were beholden to certain trial-dictated behaviours, mostly regarding concomitant medication use. However, two more dominant constructs, "Trial citizenship" and "Deferred responsibility" may explain some of the differences observed between the sites.

Trial citizenship in South African inpatient volunteers

South African participants described their important role, indeed sometimes their *work*, in facilitating the trial's success. They were largely knowledgeable of, and aligned their work as trial participants with, researchers' objectives. On hearing a participant was withdrawn due to a contra-indicated medication prescribed by her own doctor, they felt she was correct in reporting it even though she could not play her trial role anymore:

"Interviewer: So, how did you feel about [another participant being withdrawn from the trial for taking a contra-indicated medication]?

Respondent 3: We feel bad

Respondent 5: We feel very bad

Respondent 4: But [on] Saturday, the doctor told us [about] the side-effect [of] the tablets that the girl was taking for the skin. It will affect what? [Respondent 2: kidneys]. .. Then [that] girl, she can't work with us." [FGD 02 South Africa]

Even though participants conveyed this sense of responsibility towards the integrity of the trial, there were second-hand reports of others not playing their role properly, deliberately withholding information for fear of being withdrawn. Though loss of reimbursement was cited as a reason for this, one participant expressed it in relation to a worry about having to leave because participants had *"come [to] enjoy the trial"* [IDI South Africa, respondent 11]. This South African trial was conducted under stricter conditions than the trial in Tanzania in terms of, for example, admission for intensive pharmacokinetic sampling periods, and the foods and concomitant medications allowed. Being with each other and the staff for 72 hours offered a liberating experience; there was much mention of the opportunity to speak freely with others who understood what it meant to be HIV positive, phone numbers were swapped and focus group discussions were happy reunions. This may have nurtured the trial citizenship underpinning some decisions regarding relevance to report, which were sometimes made by consensus. Thus, for instance, an AE was considered unrelated to the South African trial if others were not experiencing it:

"Somebody will wake up and say 'Guys I am feeling this, is anyone feeling it?' And then because that one said no… you will also think 'Ah maybe it's me. It's only me'" [IDI South Africa, respondent 05]

Over and above these participants' roles in this trial, however, there was recognition that they were monitored to avoid personal risk. Perhaps in addition to trial citizenship, the fact that the antimalarial trial drug was being used experimentally in participants without malaria meant that they had a heightened vigilance for possible side-effects:

"*I was busy trying to look for side-effects from the trial.*" [FGD 01 South Africa respondent 4]"*They asked me how did I feel.. and what the medicine [did] to me"* [FGD 02 South Africa, respondent 3]

Deferred responsibility in sick Tanzanian outpatients

In Tanzania, particularly in the in-depth interviews, there was little discussion, or sometimes understanding, of the trial. Some participants showed how the locus of responsibility for knowing relevant information fell with trial staff as clinicians, rather than with themselves; the doctor had the knowledge to prompt them to reveal whatever additional information was required over and above what they respond to a general enquiry:

> "*So I think the doctor understands more, that's why he went on to probe. So I didn't forget or wasn't careless, but it's my knowledge that is low, that if the stomach aches then even diarrhoea may be there, 'what about diarrhoea?'*" [FGD 1 Tanzania, respondent 1]

Tanzanian participants were recruited to the trial individually when they presented to routine services with malaria symptoms for which they sought a cure. The trial was largely incidental to them achieving good health. There were few references to researchers' priorities; it was personal:

> "*I didn't just join because of the name malaria, no. But it's because there are tests and examinations that are being done on me so as to know how my health is. …So I was being researched. So I know how my state is, I was examined*" [FGD 1 Tanzania, respondent 1]

The experience predominantly described was of feeling ill, joining a project (to get optimal management), and thus feeling better. Participants considered the questions the doctor asked about health and treatments were for their personal benefit:

> "*When he asked what kind of drug have you used, I suspect that he asked so that we don't make it a habit buying drugs from drug shops, we should come and get tested first*" [FGD 1 Tanzania, respondent 4]

In the in-depth interview, responses to questions about change in health since baseline overwhelmingly concerned gratitude for improvement, malaria-related or not.

5.3.3.3 Trial clinicians' reflections on the limits of sensitivity

Participants in both sites had overwhelmingly recommended that more detailed questioning (checklists or in-depth interviews) helped them to report, and Tanzanian participants said that the focus group discussion taught them the importance of maintaining their own detailed illness and treatment record, for its personal health benefits. Trial clinicians, meanwhile, spoke of the challenge of eliciting comprehensive but relevant data when they will never know everything. For well-studied drugs, the focus of more detailed questioning could be on known or anticipated risks, combined with general enquires to detect anything else. But it was a quandary whether to probe for AEs that are perhaps insignificant or irrelevant to both clinicians and participants:

> *"If they choose not to tell me about their headache when I'm asking them how they are and how they [feel] in their body…how severe could it be or… how important [is it] to them?"* [Trial clinician 2]

For trials of drugs with known safety profiles:

> *"Isn't the grading such a key issue? Because people's lives are full of minor mishaps, and minor symptoms all the time. And if you really, really wring it out of people, you could generate a lot of irrelevant grade 1 stuff*" [Trial clinician 4].

There was, however, concern that selective detailed questioning could miss minor illness that impacts on adherence, and thus efficacy and an increased risk of malaria resistance at a population level. These clinicians, reflecting how patients may be intimidated by their role as doctors, despite them being at pains to be otherwise, suggested that other cadres of staff be involved in order to overcome barriers to reporting. This involvement, they said, could include designing elicitation strategies (social scientists), questioning participants (nurses or social scientists) or interpreting safety results (anthropologists).

5.4 Discussion

Clinical trial guidelines have been developed to attain common standards across phases of drug development and disparate sites (1). However, little is offered about methods for

questioning participants to elicit information pertaining to harm and tolerability outcomes. The issue of whether to employ checklists in trials, as is routine in other pharmacoepidemiological study designs, and, if so, their components and mode of delivery, remains contentious. This study shows, as others have, that asking participants to indicate which of a checklist of items applies to them increases the sensitivity of detecting data compared to a general enquiry. In addition, some data were only detected during subsequent in-depth interviews, suggesting that checklists overcame some, but not all, factors involved in under-reporting in response to a general enquiry. In these two trials the additional data gained by more detailed or in-depth questioning did not alter eligibility for participation, or detection of AEs which staff determined to be related to the study drug(s). However, because of the inherent limitations of individual causality assessments, lost AEs may be ADRs when analysed statistically in a pooled analysis. The findings of this qualitative study suggest that contents of a checklist cannot feasibly be exhaustive enough to trigger all issues forgotten, deemed insignificant or irrelevant. It is also unlikely that checklists alone will overcome fears of reporting, by those who perceive themselves as either participants or patients.

Data reports in context

This study offers empirical evidence of the impact on antimalarial and ARV trial safety endpoints of individual perceptions of what is (and is not) memorable, significant, relevant, of personal consequence, articulable and necessary to report. Recall was a dominant factor, reflecting the focus on memory in the broader pharmacoepidemiological literature about validity of participant-reported data (173). In addition, this study provides evidence of some less-studied influences on reporting, including participants' perceptions of significance (whereby illness or treatment use may be normalised or gauged against another), relevance to report and negative consequences of reporting. Such factors influencing participant reporting in clinical trials are similar to those observed in other areas of epidemiology, suggesting that the evidence base for elicitation could be extended by learning from relevant methodological work conducted in other disciplines (239, 240).

The study also identifies factors influencing elicitation processes and outcomes that could be more specific to the clinical trial context. Perceptions about relevance to report and fears of reporting appeared most likely to be influenced by differences between the trial

sites. Others have explored how trial participants understand their identity, as being somewhere on a continuum from patient to active volunteer, and how this may shape outcomes (195-197). This study describes how the South African participants appeared to take on a form of "trial citizenship". Their prolonged opportunity to interact with staff in the ward is likely to have affected their understanding of the trial and what was important to researchers in a way that differed to the Tanzanian outpatient participants. They viewed reimbursement as financially welcome but it could also have afforded them a responsible, job-like, role in determining the trial's outcome. Beyond this, these otherwise healthy, HIV positive South Africans described their experience in the trial as a treat, a space to be free from a complicated life. The small community of participants in the trial ward, situated in a wider context of AIDS activism and social mobility, may reflect active biological citizens in allegiance with the conventional biomedical community of clinical trial staff (241). The Tanzanian situation, meanwhile, is more likely to have represented a typical malaria therapeutic efficacy study whereby participants seek a cure from a clinic as usual and are focused on recovery. As observed elsewhere, despite consent information detailing experimental aspects of the trial, the consultations were understood by Tanzanians largely for their personal medical benefit (242).

Locus of responsibility for assessing causality

This study did not reveal whether the design and contextual differences between the two sites fully explains the relatively high number of AEs in South Africa compared to Tanzania, though it may be that there was greater awareness of the focus on detecting side-effects in the former trial, compared with the latter where the focus appeared to be on resolution of malaria symptoms. South African participants were also relatively healthy, which may have increased their reporting of any deterioration in health. If so, this signals the potential for the trial design to influence how safety end points are understood and reported. While individual causality assessments made by staff during a trial allow for clinical decisions and adherence to regulatory reporting timelines, a central tenet is that evidence for causality is determined for registration purposes on the basis of aggregated data at the end of a trial or on meta-analysis with data from similar trials. The findings of this study suggest that implicit assessments of causality by participants can occur very early in the process of data generation, undermining the assumptions behind the 'objective data' used to determine harm and tolerability. Participants make decisions to report based on their personal logical assessments of causality, severity and risk-benefit. Thus, the lines between adverse events

and suspected adverse drug reactions according to ICH definitions of causality are blurred at the point of questioning (1).

The importance of treatment literacy

There were significant differences in how ARVs were talked about by participants in these two sites. Treatment literacy - i.e. knowing about medications one is taking - is fundamental to pharmacoepidemiology, especially where record linkage is difficult. Yet appropriate ways of eliciting concomitant medication reports in trials is a particularly neglected research area. Lay public definitions of what constitutes a medication may explain observed delays in reporting some medications versus others (71), and this study's data from Tanzania suggest that drugs considered not relevant to the hospital are not reported. Others have shown how people may not know names of prescribed medications, so rather leave this knowledge to a health practitioner (243). This behaviour may be more likely in more paternalistic healthcare systems (244). However, HIV is a particularly stigmatised condition, and references to it (and its treatment) may also be subtle, even by providers. While South Africans also appeared not to know names of prescription-only drugs, their ability to intimately know ARVs could be a result of greater progress in de-stigmatising HIV/AIDS, particularly in a setting where participants are likely to be empowered by well-known advocacy groups and treatment programs (245). Though the Tanzanian respondents did not express themselves as trial citizens, their usual experience of malaria management was certainly enhanced by the specialised, dedicated trial team. By the time of their focus group discussions they spoke more knowledgeably about the trial, expressed satisfaction with the focus group discussion allowing them to socialise, dropped euphemisms for HIV, and proposed they take more responsibility for keeping health and treatment records. This suggests there is scope for altering what could be the natural course of a patient-health facility relationship to overcome specific impediments to trial reporting. However, this would need careful consideration in view of evidence that moving away from usual care to the trial context may influence how information that becomes safety end points are understood and reported (246).

Blood tests versus reports

Blood taking is a strong symbol of biomedicine, and there may be unrealistic expectations about the capabilities of blood tests in trials (242). The privileging of blood over recounted

illness and treatment experiences may reflect the recognition of a technological imperative in biomedicine; that tests and equipment have the power to provide doctors with the answers they require (247). This may be reinforced by a lack of opportunity to describe illnesses, with health workers indicating what is important to say, hear and know (248). If trial participants or patients are to be encouraged to relate their experiences of illnesses and concomitant medications, they may need reassurance that their experience is of equal importance and that blood tests may not provide the same information. Those eliciting responses may need to show particular patience in listening to stories that surround trial data end points.

5.5 Limitations

This study does not reveal a whole 'truth', as participants' recognition and willingness to report information vary, and are subject to the trial clinicians' and researchers' interpretations. Other than lack of a gold standard, a potential limitation is that participants who did not report differently between general enquiries and checklists were systematically different from those who did. As the goal was to explore reasons for non-reporting, this study focused on the latter, but future work could include the former to understand reporting more generally. Interviewing participants immediately after each visit where general and checklist enquiries were compared could have increased the validity of comparisons, but could also have led to participant fatigue, and it was opted to use time and budget judiciously, delaying in-depth interviews until after Day 7. The sample size was too small to detect statistically significant differences in responses to the 3 questioning methods and it was not possible to measure influence from other factors, such as the characteristics of the trial clinicians and interviewers (including role and gender), precise mode of delivery of questions, or influence from the participant trial information. The questioning methods were compared in a sequential design rather than in a direct parallel manner as this was a realistic representation of how these methods are generally used in trials. The study results may not be generalizable to other trials where the methods to detect safety data and context are very different.

5.6 Conclusion

This study offers more evidence that questioning methods influence the detection of clinical trial safety data. This could prove a major limitation to optimal safety assessments,

and preclude valid pooled data safety analyses. The study offers explanations for this phenomenon through the voices of participants, indicating that different trial contexts cultivated some specific conditions that had a role in mediating recognition, reporting and articulation of important variables. Based on this work, it makes sense to develop appropriate messages for participants to convey the concept that checklist items, if used, are merely examples of the level of data required, and that their answers to questions about ill health and treatment use are as important as tests and examinations. This may particularly help those believing it is the health worker alone who assumes responsibility for reports. Participants could also be counseled to quell potential fears about reporting and appropriate ways to manage prohibited concomitant medications reports without 'punishment' could be explored. For any elicitation method, conceptualisations of the purposes of elicitation, including those that trial staff hold, should be incorporated in its design to help negate potential areas of mismatch. There is a need to find optimal phraseology to ensure understanding of question terminology. In areas with low treatment literacy, pictures or samples could help staff uncover the names of concomitant medications. Pictorial methods may also be of value in enhancing communication, including the rationale for the information needed (74). The challenge in incorporating these in clinical trials is the need to be systematic whilst being locally interpretable. Despite their limitations, checklists are an obvious focus, a way forward being to explore, over and above content, their mode of application. It is here that consideration of possible influences from a trial's social context would be most useful. Clinical researchers could also consider what role their participants may take on, for example that of a patient or a trial citizen. Further methodological work should be embedded in clinical trials to investigate influences on the measurement of participant-reported endpoints (180). Understanding contextual factors influencing trial outcomes could form the basis of innovative ways to capture important safety data outcomes (249).

CHAPTER 6:

INFLUENCES ON PARTICIPANT REPORTING IN THE WORLD HEALTH ORGANISATION DRUGS EXPOSURE PREGNANCY REGISTRY; A QUALITATIVE STUDY

Elizabeth Allen[1], Melba Gomes[2], Lucy Yevoo[3], Omar Egesah[4], Christine Clerk[5], Josaphat Byamugisha[6], Anthony Mbonye[7], Edwin Were[8], Ushma Mehta[9], Lynn Atuyambe[10]

[1]Division of Clinical Pharmacology, Department of Medicine, University of Cape Town, Cape Town, South Africa.

[2]World Health Organisation, 1211 Avenue Appia, Geneva 27, Switzerland

[3]Dodowa Health Research Centre, Dodowa, Ghana

[4]Department of Anthropology, Moi University, Eldoret, Kenya

[5]School of Public Health, University of Ghana, Ghana

[6]Department of Obstetrics and Gynaecology, Makerere University, Kampala, Uganda

[7]Ministry of Health, Kampala, Uganda

[8]Department of Reproductive Health, Moi University, Eldoret, Kenya

[9]Independent Pharmacovigilance Consultant, Cape Town, South Africa

[10]Department of Community Health and Behavioural Sciences, Makerere University School of Public Health, Kampala, Uganda

Publication status:

Published: BMC Health Services Research 2014;14:525.

MG and UM conceived, designed and/or conducted the parent WHO Pregnancy Registry with input from EA, CC, and Mackensie Yore, Esperança Sevene, Jan Singlovic, Max Petzold, Viviana Mangiaterra, Elisabeth Elefant, Frank M Sullivan and Lewis B Holmes. For this qualitative component of the registry pilot study, the candidate (EA) and UM designed the research questions, EA developed study tools, conducted the analysis and wrote the paper. LMY, OE and LA refined the study tools, supervised the local social science teams, and took part in the analysis. All authors of this paper contributed to data interpretation and approved the final version.

Karen Barnes and Clare Chandler conducted an in-depth review of the study design and manuscript during its development.

Abstract

Background: The World Health Organisation has designed a pregnancy registry to investigate the effect of maternal drug use on pregnancy outcomes in resource-limited settings. In this sentinel surveillance system, detailed health and drug use data are prospectively collected from the first antenatal clinic visit until delivery. Over and above other clinical records, the registry relies on accurate participant reports about the drugs they use. Qualitative methods were incorporated into a pilot registry study during 2010 and 2011 to examine barriers to women reporting these drugs and other exposures at antenatal clinics, and how they might be overcome.

Methods: Twenty-seven focus group discussions were conducted in Ghana, Kenya and Uganda with a total of 208 women either enrolled in the registry or from its source communities. A question guide was designed to uncover the types of exposure data under- or inaccurately reported at antenatal clinics, the underlying reasons, and how women prefer to be asked questions. Transcripts were analysed thematically.

Results: Women said it was important for them to report everything they had used during pregnancy. However, they expressed reservations about revealing their consumption of traditional, over-the-counter medicines and alcohol to antenatal staff because of anticipated negative reactions. Some enrolled participants' improved relationship with registry staff facilitated information sharing and the registry tools helped overcome problems with recall and naming of medicines. Decisions about where women sought care, which influenced medicines used and antenatal clinic attendance, were influenced by pressure within and outside of the formal healthcare system to conform to conflicting behaviours. Conversations also reflected women's responsibilities for producing a healthy baby.

Conclusions: Women in this study commonly take traditional medicines in pregnancy, and to a lesser extent over-the-counter medicines and alcohol. The World Health Organisation pregnancy registry shows potential to enhance their reporting of these substances at the antenatal clinic. However, more work is needed to find optimal techniques for eliciting accurate reports, especially where the detail of constituents may never be known. It will also be important to find ways of sustaining such drug exposure surveillance systems in busy antenatal clinics.

6.1 Background

Pregnancy exposure registries are observational studies measuring the association between drugs taken during pregnancy and the pregnancy outcome (9). They make important contributions to evidence on the safety of medicines in pregnancy, particularly as this vulnerable population is rarely included in clinical trials (41). There is a specific need to establish the safety profile of life-saving therapies in countries with a high burden of diseases such as malaria, TB and HIV/AIDS, so that pregnant women are neither denied access to a safe drug nor exposed to an unsafe drug. For instance, with the recent rapid scale up in access to artemisinin-based combination therapies (ACT), effective pharmacovigilance during pregnancy is important as these drugs are contraindicated in the first trimester (15, 250). There is also much to be learnt about the safety of non-prescription medicines, including herbs, which are widely used but rarely studied (251-253).

Data about the prevalence of birth defects, including those related to substances used post-conception, are generally lacking in resource-limited settings (middle and low income countries) so the World Health Organisation (WHO) designed a registry suitable for implementation by national health authorities in these contexts for routine surveillance (43, 254). The WHO registry, which is neither disease nor drug-specific, involves the prospective collection of detailed health and exposure data from a woman's first antenatal clinic (ANC) visit until delivery (Figure 6.1). As data accumulate within countries or regions, it is expected that the risk of specific drug exposures as determinants of maternal and neonatal outcomes may be quantified. In addition, the registry is intended to help build capacity for improved maternal and neonatal care within the health care systems. This could include training ANC staff about the importance of obtaining a good treatment history from pregnant women, collecting various indicators of maternal and new-born health, and strengthening referral systems for babies born with poor health.

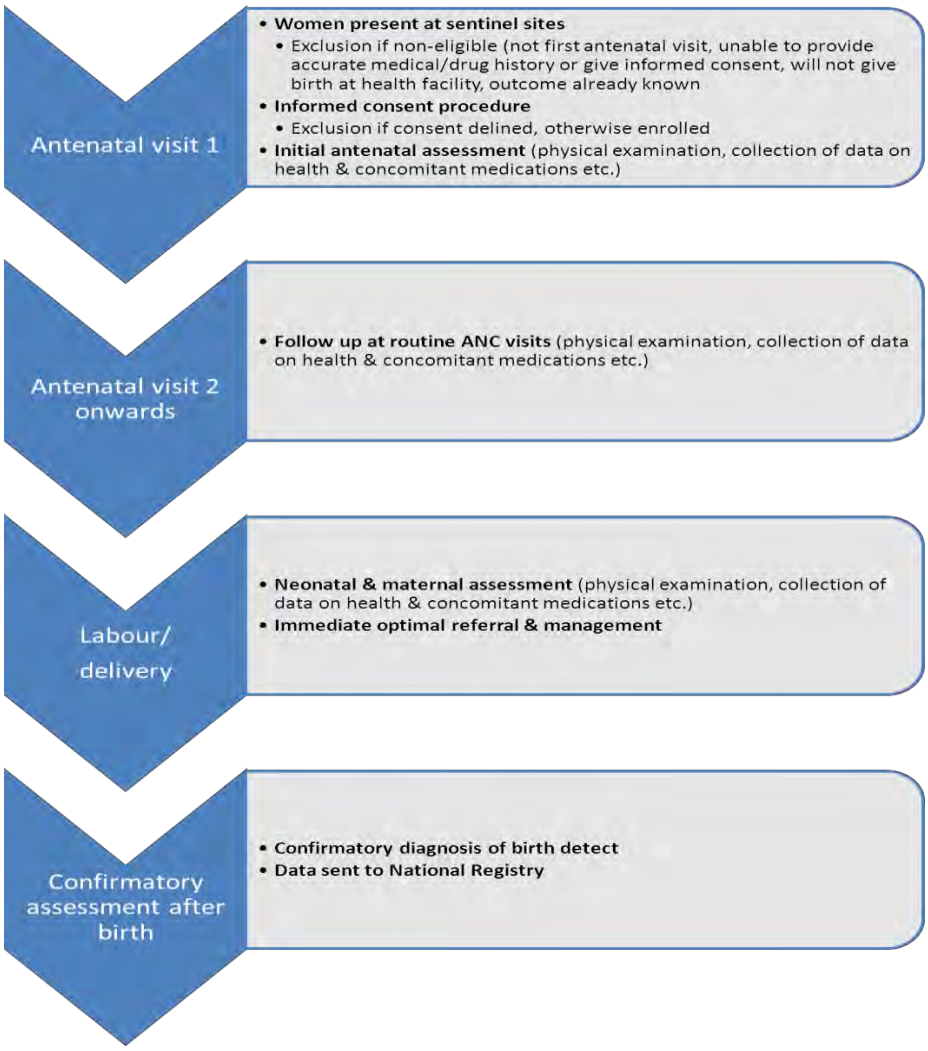

Figure 6.1: Flow diagram of WHO pregnancy registry

As with all pharmacoepidemiologic studies the success of the WHO registry relies, largely, on the accuracy of exposure, outcome and covariate data (173). In countries where there is a relatively poor infrastructure for maintaining institutional medical, prescription or dispensing records, it is difficult to derive objective measures of medical history and drug use (255). Maternal reports may be the only source of these data, particularly for medicines obtained from private clinics, shops and traditional practitioners. Methodological research shows that demographic characteristics (such as age and socio-economic indicators), recall abilities, perceptions about the significance or relevance of the information, anticipation of consequences, and the ability to articulate a response, are potential factors involved in shaping patients' or research participants' reporting of health problems and medicine use (140, 173, 256). The way questions are posed can also influence responses. For instance questionnaires that include indication-for use and drug-

specific questions increase prevalence estimates for drug use compared to open-ended questions (174). As barriers to accurate and complete reporting are likely to be context-specific it is important to investigate these in the area where research is conducted, such that relevant measures may be taken to improve data quality.

A qualitative component was included in the WHO registry pilot study in Ghana, Kenya and Uganda to examine the factors that either enabled or hindered the process of creating and maintaining a successful registry in low and middle-income settings. Specifically, it sought to explore, from pregnant women's perspectives, barriers to their reporting, and thereby staff recording, of drug use data in the registry and how these might be overcome; what types of information are under- or inaccurately reported at ANC visits, the underlying reasons, and how women would prefer to be asked questions. This paper describes the results of these research questions to inform optimal implementation of pregnancy registries by health authorities in similar contexts.

6.2 Methods

We employed a qualitative study design, using focus group discussions (FGDs) to collect the data. These were nested within a WHO registry pilot study and conducted between July 2010 and August 2011. In each registry site a sample of women attending their first ANC visit were enrolled. ANC nurses in the registry teams were given standardised training in how to elicit information on medical histories and medicine use from enrolled women using a data collection tool which complemented the existing ANC records (43). A similar tool, with additional fields related to the birth outcome, was used at the time of delivery. Women were also encouraged to maintain a written record of the names of medicines they had used (in notebooks), or collect samples or packets in plastic envelopes supplied. These were then to be brought to the site at routine ANC follow-up visits to help registry staff make an accurate record.

6.2.1 Qualitative study areas, population and sampling

The study population for this qualitative component included two main categories of women 1) women of all ages who had been enrolled in the WHO pregnancy registry pilot study being conducted at selected sites in Ghana (Dangme West District Hospital), Kenya

(Webuye District Hospital) and Uganda (Iganga District Hospital) and 2) women of child-bearing age, who had experienced at least one pregnancy, living in the communities or catchment areas from which participants for the registry were drawn. The latter category was included to potentially detect influence of the registry methods on reporting. Each site was situated in a rural or semi-rural (small town) location. National figures for women's use of an ANC for at least four visits during pregnancy were 78% in Ghana (2008), 47% in Kenya (2009) and 48% in Uganda (2011); 55% (2008), 44% (2009) and 58% (2011) of women respectively are attended to by a skilled birth attendant during delivery; for 2011 the prevalence of HIV (age 15-49) and annual number of malaria cases are 1.5% and 1,041,260 for Ghana, 6.2 and 1002805 for Kenya and 7.2% and 231,873 for Uganda respectively (257).

There were 12 strata of FGDs, as participants were stratified by country, whether enrolled in the registry or from the registry's source community, and < or ≥ 24 years old (the latter because younger women often feel less comfortable in expressing themselves in the presence of older women (258). Sampling was purposive as regards the strata and where community participants resided in relation to the respective ANCs. Those enrolled in the registry were accessed through liaison with the pilot registry ANC study staff teams, while women in the source community groups were recruited within the catchment areas of each antenatal clinic through households and community groups such as markets. Two to three FGDs were to be conducted per strata, with a minimum of six participants in each. See Supplementary File 6.1 for the question guide.

6.2.2 Study conduct, data management and analysis

Senior local social scientists (LY, OE, LMA) supervised the field conduct of FGDs by teams trained according to a manual and standard operating procedures developed by an international coordinator (EA). The original English version of the FGD question guide was translated and pre-tested in at least one FGD per site to ensure the terminology was locally appropriate. Topics included experiences and perceptions about treatment-seeking in pregnancy, poor birth outcomes, maintaining health and treatment records, the pregnancy registry and reporting information at the ANC. As recall ability is related to the time since exposure to a medicine or other substance it was also important to understand the influences on seeking adequate antenatal care, particularly the first ANC visit, and the types of medicines used. The FGDs were held close to the ANC but in a private space. Audio

recordings of each FGD were transcribed directly into English using a meaning-based method, checked for quality and imported into NVivo 9 (QSR International Pty Ltd(237)). EA and each country social scientist co-coded at least two transcripts (deductively, using the research questions, and inductively) to agree on an initial coding framework before EA completed coding for all countries. Each transcript was read several times before relevant text was examined for repeating concepts (codes) which were labelled and grouped into categories and themes reflecting the underlying meaning behind statements (211). Quotes that represented the categories and themes were then selected for inclusion in this manuscript. Code counts were also used to express the size of some categories.

6.2.3 Ethical considerations

Approval from the ethical committees or boards of the following institutions was obtained before the WHO pilot registry study started in each respective country: the WHO, Ghana Health Service, Moi Teaching and Referral Hospital, Makerere University Faculty of Medicine, and the Uganda National Council for Science and Technology. The University of Cape Town Human Research Ethics Committee also approved the analysis plan. Informed consent processes and forms for the FGDs were available in the local languages and, though information could be explained in a group, each woman met with a facilitator to confirm consent. Participants were informed about who would have access to the data collected, that refusal to participate or withdrawal of consent wouldn't affect their medical care, and that they did not need to discuss anything they were not willing to. No participant withdrew consent. Refreshments were provided during FGDs.

6.3 Results

Twenty-seven FGDs were conducted across the 3 study sites with a total of 208 women (Table 6.1). Ghana was unable to achieve two FGDs for registry participants and compensated by enrolling more women in the community FGDs. The qualitative analysis revealed three themes: "factors directly shaping reporting of exposures at the ANC (allopathic and traditional medicines, and alcohol) "," influences on formal antenatal care" and "the social context of pregnancy healthcare-related behaviour". Categories within these themes included: "importance of reporting", "negative consequences of reporting", "memory", "ability to name medicines", "special relationship with registry staff and registry-specific tools facilitates information sharing; "good use of the ANC", "poor use of

the ANC" (with their codes); "pressure from all sides to conform to certain behaviours in pregnancy and "desire, and responsibility, for having a healthy baby". These are presented diagrammatically in Figure 6.2, while differences between the strata are described below.

Table 6.1 Focus group discussion (FGD) participants in each strata

	Ghana				Kenya				Uganda			
	Registry		Community		Registry		Community		Registry		Community	
Age strata (years)	<24	≥24	<24	≥24	<24	≥24	<24	≥24	<24	≥24	<24	≥24
Number of FGDs	1	1	2	4	2	2	2	2	2	3	3	3
Number of participants per strata	9	9	16	40	13	12	20	21	11	19	19	19
Mean age (range) in years	21 (18-23)	31 (26-41)	21 (18-23)	38 (24-72)	20 (NK)	30 (NK)	23 (NK)	30 (NK)	21 (18-23)	27 (24-35)	21 (17-23)	31 (25-40)
Mean number of children	0	2	1	4	0	2	1	3	1	2	1	3

NK = not know

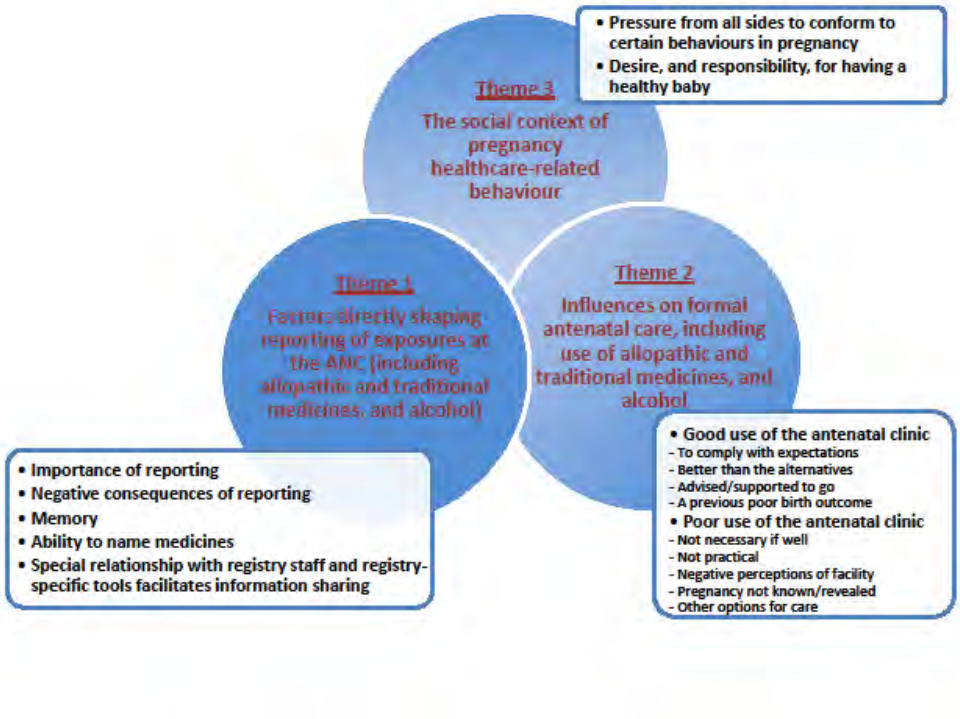

Figure 6.1 Summary of the influences on participant reporting in the WHO drugs exposure pregnancy registry

6.3.1 Factors directly shaping reporting of exposures at the ANC (allopathic and traditional medicines, and alcohol)

Two large categories within this theme (across all countries, ages and whether in the community or registry groups) were conflicting; women said that it was important to report everything relating to their health at the ANC for their own and their baby's benefit (31% of relevant codes). However, this could have a negative reaction from the health workers as they risked being scolded and sometimes humiliated in front of others (28% of relevant codes). This perceived importance, but potential negative consequences, of reporting are reflected in the quotes below:

> *"[ANC staff] are the ones who will know the medicines to give. So whether you are scolded or not you have to tell them because, when you [are] scolded and given drugs to cure yourself, you are better off*" [Older community participant, Ghana]
>
> "*Some nurses don't know how to handle patients so you just fear that she may scold you in front of people. So you just keep quiet even if you have a problem*" [Younger community participant, Kenya]
>
> " *[You have to report] the drugs you took before starting antenatal [visits] as, if you have a drug that is not good, she would tell you next time you are pregnant don't take it again. [Later]: sometimes when you say it the nurses may not be patient with you, so you are afraid that when you tell her she will be annoyed with you*" [Older enrolled participant, Ghana]

Therefore, while women said that they should (or did), report everything, they often later said that they would not (or did not). The conversations reflected a general fear of health workers' attitudes, with older women in Ghana also saying that it was easier to converse with older nurses who were more polite. When asked about substances they had used during pregnancy, women across all strata said they particularly feared reporting their use of traditional medicines, to a lesser extent over-the-counter medicines and alcohol. There were numerous descriptions of personal and others' experiences of being scolded for using these, including women being accused of using unreported traditional medicines during labour to account for a difficult birth.

There were exceptions, with sporadic descriptions of good relationships with ANC staff in general, where use of such substances was discussed. Moreover, there was limited mention in Ghana and Uganda of ANC staff endorsing the pregnant women's use of traditional medicines, and of participants' perceptions that nurses accept its benefits:

> "*I know traditional medicines work better than what the modern health workers tell us ……When you tell them they don't accept [it], yet they also use [it] and they will never tell you that they [do]*" [Younger community participant, Uganda]

Smaller categories in this theme related to memory and the ability to name medicines (12% of relevant codes). There was little explicit mention of these impacting on reporting. In fact, women in community FGDs in all countries and age groups said they do remember everything taken during pregnancy: "They remember alright but they will not say" [Older community participant, Ghana]. However, these concepts were sometimes implicit in the transcripts; women talked about substances they had used but couldn't remember or know their names and ingredients, particularly herbs.

Suggestions from women about how they could more readily report at the ANC what medicines or other substances they had used reflected the dominance of fear of reporting as a barrier. Thus they recommended that nurses be more understanding / empathetic. In Uganda, where this was discussed at length, enrolled women explained how they were able to forge a special relationship with the friendlier, dedicated registry nurses who sought, and would accept, full disclosure. Similarly, when the registry rationale and method was explained to them, community women considered that being enrolled would help them report such information. There was, however, some indication that this improved attitude of registry nurses did not always extend to those women not enrolled:

> "*Participant: You don't talk about [Waragi, a local gin] here because you may find a health worker who may be harsh on you why you took [it].*
>
> *Facilitator: What about here in the registry, would you still not talk about Waragi?*
>
> *Participant: Here I know I have my nurse who works upon me, so … if she asks me I can let her know that I take a little Waragi, not to be [a] drunkard.*

Facilitator: *What of the local herbs?*

Participant: *I also have to tell her.*

Facilitator: *Won't she shout at you?*

Participant: *No, she doesn't shout at me*".

[Younger registry participant, Uganda]

"[Registry nurses] welcome you and give you a lot of care and even ask how you are feeling. But if another person who is not [enrolled in the registry] comes they just look at her like this, yet they are the same nurses who welcomed you well"
[Older registry participant, country anonymised]

The registry also facilitated information sharing by encouraging enrolled women to collect names or examples (e.g. leaves, packets) of what they had used, with community groups also acknowledging this could be possible: "*With respect to the herbal medicine, when you can't remember the name, you will come back and pluck it for the doctor, that this [is] the drug I have taken*" [Older community participant, Ghana]. In addition, women said that making antenatal care more accessible, particularly to those in rural areas, would be beneficial for reporting:

"*I also support the home visit because the atmosphere will be conducive to reveal all your health matters, as you are relaxed. Rather than the hospital where you are in a hurry to clear the line for next patient and allow the busy nurse [to] continue with her ever busy schedule*" [Older registry participant, Kenya]

6.3.2 Influences on formal antenatal care

The analysis showed two categories related to formal antenatal care which impact on medicines used in pregnancy and the recall period for reporting; good or poor use of the ANC. The transcripts revealed that good use relied on women seeing the value of going to the ANC if, for instance, it was considered better than alternatives (e.g. traditional care givers). This was apparent in all strata and was particularly important to women who had experienced a previous poor pregnancy outcome while relying on other care options. Some women had practical (including financial) or moral support from family or friends to attend.

However, while there was limited evidence of pregnancy testing and ultrasounds being obtained at private facilities early in pregnancy in all countries, it was common in all strata to delay attendance at the government facility until the second or third trimester. Many women said they delayed attending the ANC until they experienced illness, while others intended to go only once to collect an ANC card/book. This was essential in case they needed to use the labour ward, as they anticipated abuse for not registering if there were problems during delivery. Poor attendance was also due to negative perceptions of the ANC (an expectation of being ill-treated by staff, sitting for a long time in a queue or having to undergo an HIV test that they did not want). The narratives showed that younger women were particularly vulnerable to inadequate formal ANC care. They were more likely to have not recognised the pregnancy signs and symptoms, have had less access to support, may have hidden the pregnancy initially, or attempted abortion.

Traditional birth attendants were consulted in all three countries by both those who did and did not attend the ANC regularly, particularly for positioning the baby in the womb. Other options for healthcare during pregnancy included drug shops, traditional healers or herbalists (though these were all less common as they were not considered pregnancy experts), family or community members, and spiritualists. Traditional medicines were widely used across all strata (58% of codes relating to substance-use in pregnancy). Ingested traditional medicines, including leaves and roots sometimes mixed in clay or soil, were taken for fever/malaria, constipation, heartburn, to treat or provoke vomiting (the latter to clean dirt or illness from the mother or baby), anaemia, to prepare for birth (by cleaning the womb of sexually transmitted diseases or for softening the pelvic bones and position the baby correctly), to prevent miscarriage, for the mother's energy, for growth/strength of the baby, or as an abortificant. Some of these medicines were named (typically those herbs picked by oneself or a relative), while others were not (typically concoctions prescribed by a healer). There was a particular discussion among older women in Ghana about not liking medicines from the ANC (such as iron and antimalarials), with one woman saying that, even if she attended the ANC to prevent conflict or condemnation, she did not trust such medicines:

> "*I will go to the health facility alright. If I don't go you will tell me [off]. So I will go and when I am given the drugs I will use them as pillows (put them aside). It's the traditional ones that I take, and [I] am strong too*" [Older community participant, Ghana]

Participants in various FGDs reported using the following allopathic medicines during their pregnancies: ibuprofen, aspirin, indomethacin, antimalarials, penicillin and amoxicillin. Some of these were inadvertently taken before the pregnancy diagnosis. For antimalarials this early exposure could reflect conflated narratives about fever, malaria and pregnancy signs/symptoms apparent in all strata:

> "*I used to miss my periods for even three months, so when it happened I thought it was just malaria. I bought some tablets then came to hospital. It is then that they confirmed that I was pregnant*" [Younger community participant, Kenya]

Women rarely mentioned traditional medicines as the cause of poor birth outcomes apart from those specifically used for attempted abortions. Poor outcomes were, however, said to be predominantly due to spiritual attacks (20% of relevant codes), but also the use of over-the-counter drugs (5%), prescribed contraceptives (10%) or being passed directly from parent to child (16%). There was confusion about alcohol throughout all strata. Some women felt that drinking it could harm the mother or baby, but there was evidence of others endorsing it, generally in small quantities, to clean the womb or baby of semen, infection or ingested substances (such as clay), and for ensuring the baby had nice whites of the eyes. Its use was even occasionally sanctioned by ANC staff. Alcohol was also used for enhancing women's appetites or sleep, and for easing labour:

> "*People say that Akpeteshie (a local gin), is not good. Well I confess I drink Akpeteshie when I am pregnant because they say that, if you are pregnant and you drink it, your baby will be clean. If you are pregnant and you drink alcohol, you do not get drunk, no matter the quantity you drink*" [Younger community participant, Ghana]

> "*Participant:* *You should take Guinness (a strong beer)*
>
> *Facilitator:* *Why Guinness?*
>
> *Participant:* *They say when you take it during delivery it cleans the womb, so at delivery there is no stomach ache, the womb is washed clean.*
>
> *Facilitator:* *How much do you take?*

Participant: *You take a glass. You take towards delivery, or even during pregnancy, but just once in a while you take the small Guinness.*

Facilitator: *Who says this?*

Participant: *Just people. But not the doctors of course, but women who have delivered previously and have actually used it to avoid the pain at delivery*".[Older registry participant, older, Kenya]

6.3.3 Social context of pregnancy healthcare-related behaviour

We observed two categories relating to societal influences on women's healthcare-related behaviour during pregnancy; an apparent pressure to conform to what was expected by the ANC, the family and community, and responsibility for a healthy baby. Underpinning the themes reported above was an explicit or implicit understanding that women were expected by the government ANC to attend ANC regularly, to largely abstain from self-medicating, and to report their health and treatment use accurately. Meanwhile, women in all strata showed how they may also have to adhere to certain culturally-dictated behaviours to ensure a healthy pregnancy, and manage the pregnancy at home. The latter was compounded by problems accessing formal care. Expectations of, or advice from, the different parties involved were often conflicting, and involved power relationships. For instance, the conversations with younger women revealed that they were sometimes pressurised by older women in the community to behave a certain way or disregard institutional recommendations for formal care. Over and above their own desires for a healthy baby, conversations also reflected women's responsibilities for producing babies acceptable to the father, his family, and/or the community. All FGDs showed how women must be careful about how they conduct themselves in public, what they eat or what medicines they use, as these may dictate whether a baby is born healthy. Despite these considerations, she may still be blamed for producing a child with a birth defect. These societal pressure categories are demonstrated below:

"*The old women … advised me that it is not necessary to go early [to the ANC] because you will be going there all the time. I told them that I want to go to the health centre because I was feeling lower abdominal pain. Still they rejected me and could tell my husband not to give me money for going there because my pregnancy was still young. So I just came when he gave me money to eat*" [Younger registry participant, Uganda. Attended ANC first at 5 months]

"*In my hometown, they say pregnant women should not engage in a lot of things, so if you turn a deaf ear to what the tradition says and there [are] gods in your town, [you] could have a child with a defect. Therefore if you live in such a town you should be very careful until you deliver*" [Older community participant, Ghana]

"*I also have a sister-in-law who gave birth to a child with a big head but the husband chased her and told her that he doesn't produce babies with big heads*" [Older registry participant, Uganda]

"*When one is pregnant you don't have to take alcoholic drink because it affects the babies in the womb and when you give birth the babies will not look attractive and people would not give you a helping hand in carrying him or her*" [Community participant, older, Ghana]

6.4 Discussion

This qualitative study found that women desired, and felt responsible for, carrying a healthy baby to term. However, despite largely deeming all health-related information during pregnancy important to report at the ANC to help ensure a good outcome, they could not always recall or give exact details of what substances they had used. Moreover, the pressure to conform to both formal and traditional pregnancy care options, or a practical need to combine such care, led to a common use of substances with unknown effects in pregnancy, including traditional medicines, over-the-counter medicines and alcohol. Such substances were not always reported easily at the ANC as women generally feared being admonished by staff. While pregnancy registries could play a valuable role in building knowledge about the safety of medicines used in pregnancy in resource-limited areas, they are dependent on valid and reliable data. If there is significant measurement error of exposures, outcomes or confounders, associations will be incorrect and pregnant women will not get optimal advice about suitable medicines. It is, therefore, important to consider all these barriers to accurate participant reports so that socially-contextualised solutions may be sought.

This present finding that various factors influence women's willingness or ability to seek antenatal care is consistent with results of previous studies (259-261). As well as increasing the risk for complications, those not adhering to national recommendations for such care may be less likely to receive appropriate advice about safe substance use. This will be exacerbated should women not recognise, or choose to hide their pregnancy, or lack family or community support to seek appropriate healthcare, particularly in younger women (262, 263). However some women indicated that the advice they receive about substance use in pregnancy could be inconsistent; midwives sometimes endorsed the traditional medicines and alcohol that women largely expected to be vilified for reporting. This confusing situation for pregnant women probably reflects the staff's own dilemmas with interpreting mixed messages when living and working in a context where traditional medicine is an embedded practice and sometimes the only readily-available option.

The registry fostered a relationship between pregnant women and ANC staff where some women said they were more able to report everything they had used. Indeed, preliminary data from the registry database shows that use of herbal medicines and alcohol were reported in all countries. However, this qualitative study cannot confirm whether the extent of such reporting is higher than in non-registry settings or reflects the true experience of enrolled women. If the registry did enhance reporting, this may have been influenced by other factors. For instance, the informed consent process explained why it was important to report everything, and the data collection tool required ANC staff to explicitly ask participants about their use of traditional and herbal medicines and alcohol, with space to record these details. The latter is in contrast to most standard ANC records or booklets in these countries.

There was an indication that an improved relationship between patients and staff may not extend to other women attending the ANC who were not enrolled in the registry. Enrolling all women at a particular facility in a drug exposure surveillance system would ensure that the community does not become resentful of the project or those included. However, whether it is feasible for any woman presenting at a busy ANC to receive the same attention, and feel free to report her experiences accurately, needs further investigation. As a minimum, registry sites should support the various interventions to encourage respectful maternity care, including improved working conditions for ANC staff (264). Staff

should also be guided in how to collect data on substances whose use they are typically trained to discourage. Emphasis should be placed on conveying messages about known risks or the fact that there are limited safety data for some substances, so that pregnant women understand the rationale for discouraging their use (265-267). Much could be learnt from research on smoking in pregnancy, to encourage transparency and trust in ANC communications about potentially risky behaviour (268).

Poor ANC attendance will undoubtedly impact negatively on registry participants remembering their health and treatment experiences accurately. There is a dearth of research relating to the validity of pregnant women's recall of non-prescription drugs in general, and very little about self-reported medical histories and treatment use in resource-limited settings. However, it is clear that some questioning techniques are better than others (240). For instance, structured questions indicating specific health conditions and names of treatments are likely to increase the sensitivity of their detection compared with open questions (173, 174). As detailed questionnaires place a burden on staff, the limited questions incorporated into the WHO registry pilot study reflected its intention to be adopted by busy ANCs. This may have led to under-reporting of drug exposures and developing an optimal questioning method in these contexts requires further work. Antenatal records could also be designed to encourage elicitation and recording of drug exposures by staff as a routine practice. Pictures of typical health complaints and medicines may be needed in areas with low literacy, and registry teams could learn from other pharmacoepidemiologic interview techniques, whereby participants are helped to identify the timing and duration of medicine use (269). These may be particularly useful in areas where women present late in pregnancy for their first ANC visit.

Women enrolled in the registry were happy to keep records or collect samples of what they had taken. However, the constituents and doses of herbal remedies, such as concoctions dispensed by traditional practitioners, are often unknown by consumers (251). It is theoretically possible that this information may even deliberately be withheld for fear that consumers may then source their own herbs. This problem of not knowing ingredients of such substances is supported by preliminary registry findings where these details were not always obtained. It may therefore only be feasible to capture a binary yes/no for herbal use in registries, to identify signals of potential harm for later, more specific, studies. However,

certain concoctions are well known locally to be used in pregnancy and may represent a semi-standardised regimen in some regions. If these were adequately recorded at the ANC it may be possible to establish any associated risk of harm. For allopathic medicines, registry sites could pursue intra- and inter-institutional record linkages (with other clinics, outpatient departments, labour and neonatal wards) and opportunities for innovative collaborations with other local healthcare providers, including local private clinics and traditional birth attendants. There may be scope for participant referrals or data sharing, and exploring local understanding of substances that are typically used or prescribed.

This study was concerned with the complex field of obtaining accurate reports from registry participants about their exposures to medicines and other substances. Finding appropriate questioning methods in pharmacoepidemiological studies is methodologically challenging. For instance, these participants voiced their fears of mentioning traditional medicines, and yet they also considered them as having less potential for causing poor pregnancy outcomes compared to other perceived determinants, such as spiritual attacks. The latter may reflect a global public perspective that complementary medicines are generally safe (270). Judgements made by participants about the significance of a health event or treatment, or its relevance to the situation, can impact on reporting (140, 173), and substances considered by consumers as safe may potentially be deemed irrelevant to report. Due to the lack of a gold standard for participant reported data in pharmacepidemiology, it is therefore important that qualitative methods be incorporated into any future work, to contribute to understanding the nuances involved.

6.5 Limitations

We could not verify that the reports of women taking part in these FGDs are accurate. Although the questions in the FGD guide were applied in all FGDs, there were differences in how facilitators probed the answers. Therefore some FGDs generated richer data which may impact on the findings. Our results may not be generalizable to other studies where the methods and context are very different. Ghana was unable to achieve two FGDs in the registry participants' strata, and there was an error during recruitment whereby some women over the age of 51 were enrolled into some of the four community FGDs. As the Ghanaian FGDs had the richest data it is likely these issues did not compromise the comparisons between sites.

6.6 Conclusion

There was widespread reported use of traditional medicines in pregnancy in this study's population, and, to a lesser extent, use of over-the-counter medicines and alcohol. The qualitative findings suggest that the WHO registry methods could address some barriers to reporting these at the ANC as it gave women tools to help overcome problems with recall and naming of substances, and sometimes enhanced their relationship with ANC staff. Questioning techniques for women attending busy ANCs where registries are implemented should be explored to identify optimal approaches, especially for recording substances such as herbal concoctions where the detail of constituents may never be known.

CHAPTER 7: DISCUSSION

Optimal drug use depends on a comprehensive assessment of both benefits and harms derived from clinical trials and subsequent pharmacoepidemiogy studies. Clinical trials and pregnancy registries are relatively expensive pharmacovigilance options, due to their prospective design, compared to other pharmacovigilance options such as spontaneous reporting mechanisms and database studies. However they are key contributions to the range of methods used to assess drug safety. While considerable work in many therapeutic areas has been devoted to ensuring valid, reliable and comparable measures of drug efficacy, equivalent work has not been undertaken for drug safety outcomes. This thesis aimed to examine one facet of the latter, the factors influencing participant-reported data required for safety and tolerability assessments. A better understanding of this topic will contribute to improving the measurement of relevant endpoints, which in turn will enhance the conduct and interpretation of the results of individual studies and pooled analyses. This is especially important for studies conducted in resource-poor areas, including vulnerable groups such as pregnant women, as these populations are often excluded from clinical drug development plans. There may also be poorly developed pharmacovigilance regulatory capacity once drugs are distributed on a wide scale (250, 271). Thus people living in such regions are particularly vulnerable to the under-detection of drug-related harm and important public health drug interventions may be undermined.

Four research projects were conducted to address this issue. A mixed-method study was nested in antimalarial and antiretroviral drug interaction clinical trials conducted in South Africa and Tanzania to understand from the participants' perspectives why they reported medical histories, AEs and non-study medicines differently when asked about their health and treatment-use by three different questioning methods. The parent clinical trials shared some characteristics but differed in others. This meant that observations could be made about whether their contextual differences impacted the results. The outputs of this study were included in a Cochrane systematic review of similar experiments to consider whether elicitation methods in a myriad different populations and study designs had common findings relating to the number and nature of the data elicited. As comparisons between studies and pooled analyses of data from multiple clinical trials will enable us to understand more about the safety profile of antimalarial drugs, it was important to also understand the range of methods used in this field, and why. An online survey was

conducted to uncover this information because the detail of methods used to elicit these data is usually not reported in sufficient detail in publications. Insight was also gained into the reporting of similar endpoints in studies involving pregnant women; a population at higher risk of malaria exposed to antimalarial drugs (and many other medicines), that yet is typically not included in clinical trials. This was achieved by qualitatively exploring reporting behaviour in a population of African women enrolled within, or in the source community for, a pilot prospective pregnancy registry. I will now summarise the projects' results and make recommendations for current research practices and further work.

7.1 Summary of findings

7.1.1 Impact of question type

The mixed-method study described in Chapter 5 showed that the type of questioning was an important factor in determining both the quantity and nature of medical history, AE and non-study medication data reported by participants in these antimalarial-antiretroviral interaction trials. The checklists and interviews helped participants remember some information that they had forgotten about when asked a more general question about their health and use of treatments. The more detailed questioning also signalled the need to report AEs or non-study medications that were not particularly significant to them or that they had not considered relevant. These findings are consistent with what is known about the use of specific questioning to uncover pharmacoepidemiological data (173). The descriptive survey in Chapter 4 revealed a range of methods being used in various permutations by malaria drug researchers to elicit AE, medical history and non-study medication reports. Overlapping rationales for how they were chosen and implemented may reflect polarised opinions, as well as a lack of evidence for which methods are most appropriate to elicit AEs, as was described in Chapter 2 and shown in the systematic review (Chapter 3). On the one hand there is concern about under-reporting if simple general enquiries are used, while on the other there is concern about increasing the burden of more detailed elicitation methods which may only be collecting insignificant or irrelevant data. More worryingly for a researcher, however, is the potential that AEs may be evoked by more specific questioning, thus raising the probability for false data reports. Concerns about underreporting are validated by the findings in Chapters 3 and 5; more detailed questioning does increase the sensitivity of detecting AEs that are missed by more general questioning. This phenomenon is consistent regardless of the study design, population, or therapeutic area.

As regards inducing of AE reports, a nocebo effect has been postulated to be related to expectation of ADRs, prior experiences of ADRs or perceived sensitivity to medicines (70, 81, 236). Leading questions are always inadvisable; however, it is unclear whether asking clinical trial participants to indicate which symptom they experienced from a list of options in a neutral non-leading manner has a nocebo effect. Asking questions about symptoms may not influence reporting in the same way that they can, for instance, influence behaviour (272). The mixed-method study reported in Chapter 5 and the systematic review do provide some evidence, however, that the nature of AE reports elicited, in terms of their clinical relevance, severity or perceived impact on participants, may be related to the question techniques. Unfortunately, the heterogeneity of the studies in the systematic review precludes drawing firm conclusions about this issue.

7.1.2 The influence of the trial context

Participants' understandings about which information was relevant to report manifested differently in the study sites reported in Chapter 5; South African healthy HIV-infected participants had the trial in mind when considering formulating a reply, while Tanzanian malaria/HIV participants were focussed on their malaria episode. There is other evidence of the trial context influencing how participants perceive their role and them struggling to differentiate between research and medical care (195-197, 242, 246). However, results from this study conducted as part of the thesis show that the context can also influence the measurement of safety end points. This was also evident for a fourth theme, the consequences of reporting. South African 'healthy' volunteers sometimes deliberately withheld information about ill health or use of non-study medicines in case it jeopardised their participation in the trial, while Tanzanian patients withheld information about their use of traditional medicines as they perceived it may jeopardise their medical care within the hospital. These apprehensions were not generally overcome by the different questioning methods applied in either trial. They were either expressed during interviews as second hand reports about others withholding information, or revealed in focus group discussions only once participants had finished their trial participation.

The women associated with the WHO pregnancy registry also expressed concerns about revealing their use of traditional medicines as it was likely to provoke a disrespectful

demeanour towards them by ANC staff. As for the Tanzanian malaria patients above, this is a mismatch between the substances that women use and what they feel they can tell the formal medical establishment they use (273). Capturing these experiences is important as they provide valuable information on whether such substances are likely responsible for morbidity during a pregnancy, such as from drug-drug or drug-herb interactions. Moreover, if a participant in a pharmacoepidemiology study reports taking a non-study drug or other substance, it allows for a discussion about a health event (and thus possible AE) that may not already have been mentioned. There was evidence, though, that the special case of being enrolled in the WHO project sometimes overcame these problems.

That this same issue was not observed in the Tanzanian clinical trials requires more investigation. It may be that the registry participants had a longer time to build up a relationship with their nurses over their pregnancy, something inherently different between these clinical research designs or populations, or the way the methodological studies were implemented. Pregnancy is, however, a particularly vulnerable time when women in these regions are reliant on both the biomedical and traditional medical paradigms for various cultural or structural reasons (274, 275). Nurses themselves are also positioned within both zones, which may explain the evidence from this project of a tacit understanding by pregnant women that nurses sometimes recommend the use of (or themselves use) traditional medicines and alcohol in pregnancy. This may have played a role in enabling registry participants the agency to report what they would normally not have been able to say in the hospital setting, and for the nurses to accept this, once both parties had established a special and more open relationship within the safe space of this project.

It was proposed in Chapter 5 that the South Africa participants took on a 'trial citizenship', working together to contribute to the success of the trial. Kelly and Geissler have suggested a similar 'experimental citizenship' for a tentative and temporary collective of participants in an HIV clinical trial in Africa based on 'mutual claims and responsibilities' (276). The value of that trial's transport reimbursement was considered a key component of the exchanges between participant and researchers. These included information, samples etc. and being 'good citizens' of the trial on the part of participants, and medical and other care afforded by the staff and trial infrastructure. The reimbursement was understood as a payment by participants and staff alike, though this was not discussed openly. The value to participants taking part in the South African trial reported in this thesis could, at face value, be limited

to the reimbursements they received as there was no personal medical benefit for them in taking the trial medicine. However, it is possible that the trial citizenship (responsibility) that they embodied is linked, for instance, to a biological citizenship of being HIV positive (241). Thus, coming together with a shared illness in an intimate space over many days and nights, where staff had time to engage with them about this illness and its impact on their health and welfare, was of value over and above monetary gains of being enrolled in the trial. Each trial's particular context is therefore important to consider when thinking about potential impact on data end points.

7.1.3 The wider, structural context

Other areas where questioning per se (however questions are phrased) may not be enough to overcome barriers to reporting is the inability of participants to know which other medicines they had taken, and their tacit belief that what they report is less important than what a test result can contribute. The findings relating to these two themes (and the presumed consequences of reporting) suggest that other novel and context-specific approaches are required to enhance reporting. Such approaches may, however, need to reach beyond the immediate trial environment, as impediments to successful communication between patients and healthcare providers may be situated at an institutional or even community level. Poor literacy and health literacy are likely to be related to socioeconomic factors such as lack of access to schooling, information about health and the use of medicines. It may also be linked to power-relations between communities and the medical profession or governmental institutions. In the South African trial, the HIV infected, but otherwise healthy, participants were located somewhere between a completely healthy volunteer (as would be typically enrolled in a phase 1 trial) and a malaria patient who would obtain some benefit from the trial intervention. The dynamic between research and medical care, a key consideration of the Declaration of Helsinki, may at first glance be deemed less relevant for healthy volunteers (277). Yet such trials do not merely involve a researcher-participant relationship rather than a doctor-patient relationship. In any culture, perceptions of the medical profession will be underpinned by societal influences, such as their members' relative status/hierarchy. Meanwhile, clinics and hospitals have their own institutional social meanings (247). More specifically in clinical trials, great attention is given to the wellbeing of participants regardless of whether they are a healthy volunteer or patient; they are all closely monitored and the trial team is intimately involved in quickly facilitating adequate care for

inter-current illnesses or adverse events. Thus, the healthy research participant is also subject to a medically-orientated relationship with healthcare professionals. It is therefore important not to make judgements about how the trial design, and the wider structural context, may influence data reports without exploring these issues from the participants' perspectives.

7.1.4 The importance of considering social relations when designing tools for data safety collection

Problems with being able to recall or articulate the names of medicines were found in the both the clinical trials in South Africa and Tanzania (Chapter 5) and the pregnancy registry conducted in Ghana, Kenya and Uganda (Chapter 6). The practical registry tools, books and envelopes in which to collect written notes or packets of medicines used were, however, said by women to help overcome these problems. Similar tools were not employed in the mixed-method study reported in Chapter 5 and there was a death of tools reported in the papers synthesized in Chapter 3 or mentioned by survey participants in Chapter 4. However, the development of novel, practical tools for eliciting health research information is an established and growing field. Participant-held diaries have been used in economic studies, and to capture diet, sleep behaviour, aspects of lifestyle such as risky behaviour, and experiences of health, including in low income settings (278). However, there are concerns about their use in areas of low literacy. Innovative ways to overcome such problems include using pictures as a tool for a field worker to use as a conversational aid (a type of prop) for eliciting AE reports (74, 75, 278). It is important to develop such tools in close consultation with the staff and participants in order to be culturally relevant and sensitive (278). This is essential also for exploring the use of other tools during study visits to aid elicitation of non-study medication reports (e.g. samples of medicines and herbs) or mobile phones.

Aside from the technical and practical aspects of developing such tools, it is important to reflect that, despite some different design characteristics, clinical trials and pregnancy registries have a commonality in that they rely on social interactions in order to collect subjective data. As seen above, some meso (trial) or macro (structural) level contexts invoke concern in participants for reporting particular data which the questioning method alone may not be able to overcome. There may also be influences at a micro level, relating

to personal issues such as gender, language or other differences between the actors involved, personal beliefs about medicines, or response-shift (79, 231-233). Many of these issues will impact the one-to-one communications necessary during trial consultations. It is therefore important that, when developing questioning approaches, staff involved in such consultations is adequately prepared with the necessary skills to develop relationships that will allow for open communication (73). One approach proposed for collecting AE data is to remove the participant-staff interface entirely, in order to better represent the participants' experiences (119). It is, however, as yet unclear as to whether this is an optimal approach for any or all contexts, and how such data will be then used (whether in place of traditional AE listings, as a separate dataset to be investigated alongside traditional AE listings or as a preliminary step to inform a traditional face-to-face consultation) (168). There is clearly a need for more exploration of these topics in drug safety research in general and also in resource-poor settings.

7.1.5 Strengths and limitations

This thesis is particularly strengthened by its use of different methods to address the aim and objectives. A systematic synthesis of studies comparing clinical trial AE elicitation methods has identified some common findings and gaps for future work. The review itself did not identify any such comparisons in a resource-poor setting, validating the value of the work nested within South African and Tanzanian trials conducted as part of the thesis. None of the previous methodological work in clinical trials has asked participants as to their opinion about their reports of health and medicine use when asked by different questioning methods. This is a missed opportunity to build empirical evidence for the hypothesised barriers. Inviting participants to help unravel the interaction between them and the trial staff in this work has provided such evidence, and also revealed that some barriers to reporting are context-specific. This study and the qualitative work nested in the WHO pregnancy registry have generated valuable data from five African countries about the factors that influence the reporting of important pharmacoepidemiology data. Finally, as elicitation methods are rarely described in publications, it was of value to survey antimalarial drug researchers about the methods they use to elicit these data and why in order to pave the way for future conversations about harmonising appropriate questioning methods.

The research was also strengthened by rigorous conduct. The mixed method study was conducted within a clinical trial environment by experienced team members familiar with the recording of these data endpoints. The qualitative component for this study and the nested pregnancy registry work used the same operational framework which was specifically designed for health research (279). This meant that there was common training, standard operating procedures and associated tools, social science oversight, coding of transcripts, analysis and review. The survey was conducted within a sophisticated digital tool with secure automatic download of data extracts. The systematic review protocol was developed using Cochrane software and subject to independent peer review by the Cochrane Adverse Effects Method Group. The data for the systematic review were also either assessed independently or subject to a 100% review.

There are several potential limitations that have been stated in each study's chapter; the most important ones will be mentioned here. The most obvious limitation of this and similar research from a pharmacoepidemiological, positivist, viewpoint, is the lack of a gold standard for much of the data reported by participants. Validity, however, was not considered a measurable/achievable objective for the work nested within the antimalarial/antiretroviral interactions trials and the pregnancy registry. Where data are collected by a healthcare worker there is also the potential for them to interpret the response or make an error in recording it in a case record form rather than record verbatim. The studies chosen for this thesis were informed by the opportunities available at the candidate's workplace, and in Chapters 5 and 6 by what was possible to embed in important, ongoing clinical research projects with their own aims and objectives. What have emerged are some cross-cutting findings relevant for both clinical trials and pregnancy registries. As seen above, the collection of safety data, as with most exchanges of information, relies on communication and social interactions between parties that is likely to be influenced by context at various levels. The mixed-method study in Chapter 5 included a limited discussion with investigators about this topic and the survey in Chapter 4 elicited opinion on optimal methods for capturing these data. Including the staff's perspectives about these specific data may have allowed for a better understanding of the social interactions. However, data were subjected to standard internal quality control procedures for the South African and Tanzanian clinical trials and the pregnancy registry. Common themes revealed in both these studies may suggest that the findings are robust.

The studies reported in Chapters 5 and 6 used different methods and, as such, each is informative on its own terms. However this may have impacted the direct comparability of their results; the women in the pregnancy registry were not offered the option to report their experiences through two questioning methods and help explain any differences. I therefore cannot conclude that considerations about significance and relevance to report AEs, past medical history or use of medicines that were not revealed in the pregnancy registry study are not factors of this environment. However, fear of the consequences of reporting is the dominant factor for pregnant women as women mentioned this most often, signifying that this is a particularly strong concern of theirs.

Neither of the malaria drug interaction trials reported in Chapter 5 enrolled children, who suffer a disproportionate burden of malaria and whose safety outcome data are collected largely through a caregiver prox. It is therefore beyond the scope of the thesis to address in any detail the specific issues around collecting safety assessment data in triadic communications (280). I also did not address the issue of confounding by indication which is where AEs can mirror the symptoms of the disease being studied (92). This is important for malaria, but requires different methodology in order to understand this issue. Finally, while the outputs of this thesis are intended to have a broad reach within the malaria drug research community and for drug safety assessment methods in general, the studies themselves are limited in scope as regards populations and methods and thus may not be generalised to other situations with different contexts.

7.2 Recommendations

Based on this thesis' findings the following recommendations are made:

- Regardless of which method is used to elicit participant-reported data contributing to a safety assessment, it is important that staff at all levels of data collection understand the rationale for the choice of method and how to apply it in a standard manner as much as possible. Staff must also be given training in developing the skills necessary to develop relationships with participants that support open communication.
- Drug research communities of practice, both in general and field-specific, should work towards harmonisation of how safety data are collected, assessed and reported, as is being done for efficacy end points (198).

- There should be more methodological work about this topic as there is a notable difference in the attention that has been paid to the concept of validity of participant-reported data in pharmacoepidemiology research (although little for pregnancy registries specifically) compared to clinical trials, despite the latter being a relatively more regulated area. Likewise, there is a need for more investigation as to how the social interactions and contexts at the individual, trial and structural levels impact on the elicitation of these data. This may be achieved by extending community engagement activities.
- Regulatory authorities and others who guide researchers should emphasise more how different elicitation methods have the potential to impact on study results. Researchers also need clarity on how to incorporate AE data reports obtained directly from participants without interpretation; should data be submitted alongside AE listings which reflect clinician's interpretations of the reports, or should they be used during the study visit to inform the latter.

Ideally clinical research staff should be involved in the design of the process and materials for eliciting participant reports, and their training and capacity-development should include role-play, observation and feedback until staff are familiar with enacting the method in practice (74). It will be critical to hear from healthcare workers 'co-opted' from their usual job into clinical research projects about their own concerns with obtaining the required participant-reported data. Implementation of data collection should then be monitored, with re-training given where necessary (281). The approach should be described in any ensuing publications so that readers can take this into consideration when interpreting results (5). It is important at all stages of data collection and reporting to be clear about definitions and terminology for AEs (and other important variables), the frequency time period and over which they were ascertained, and how they were graded, assessed for relationship to the study drug, coded and tabulated/reported (127, 218, 224). These are immediately achievable improvements that could be made when collecting safety data.

Successful collaborations, indeed communities of practice, such as those reviewed in psychiatry, oncology and rheumatology, schedule periodic conversations to define priority outcomes of interest, and assign working groups to conduct methodological research to inform the choice and application of outcome measures; the outputs are then put to the wider research community within that field to achieve consensus about a final product (63,

201, 282, 283). We have started this process for the assessment of safety in uncomplicated antimalarial clinical trials through a modified Delphi technique study which is ongoing. Antimalarial researchers, predominantly those sensitised to this topic through the survey reported in Chapter 4, were given a summary of relevant literature and asked to rate which elicitation methods they considered relevant, important and feasible for collecting participant-reported AE and non-study medicines. Once this Delphi is complete we will invite interested parties from industry, regulatory authorities and academia to agree on a plan for priority areas for further work relating to this topic. This could entail nesting methodological research within studies to explore different approaches or tools (such as the use of pictures, props or mobile phones) for questioning participants to obtain AEs and related data (see below). This could inform whether to adopt an existing toxicity grading scheme from another therapeutic area or develop one for malaria endemic populations, or adopting a common causality assessment tool. For the study of drugs in pregnancy in resource-poor regions, the WHO is in the process of engaging with others who have conducted similar research to coordinate implementation of a prospective registry. There will therefore be the opportunity to also standardise data elicitation methods for this work. Common elements for harmonised approaches could be based on what is known about important drug effects and potential barriers to reporting. However, there will need to be scope for flexibility reflecting the outputs of any formative work within a community where research is situated, if available. While this means that a method may not be completely standardised across all research sites, there would be a core of commonality to strengthen interpretation, comparisons and pooling of data from multiple studies. Strategies should be piloted within each study context according to a pre-defined plan with observation and interviews to assess understandings and interpretations on the part of participants and staff.

Methodological research in pharmacoepidemiology is hampered by lack of an adequate gold standard, particularly for the extensive tools necessary for a general AE assessment which has numerous, often unpredictable, potential end points. Improving record linkages, using blood samples for pharmacokinetic analysis of commonly used medications and comparing new strategies against existing ones may be options. These are more challenging for resource-poor regions and/or costly, or would only offer a measure of concordance (25, 173, 255, 284). However much could be learnt from how PROs are developed as these instruments incorporate qualitative methods to understand and represent the participant's subjective viewpoint while still allowing for standardisation in

the way information is elicited (285-287). Should researchers agree that there is value in limiting data collected in terms of, for example, clinical relevance, severity and bother, then both quantitative and qualitative research will be invaluable (288, 289). In addition, we should use the social sciences to explore the extension of community engagement in clinical research to contribute to our understandings of how participant-reported outcomes are obtained and understood in low resource settings. It will be important to negotiate the tensions between the positivist pharmacoepidemiological perspective and other perspectives that, for instance, consider that the health experience of an individual is constructed by humans and may depend on social, historical or other contexts (177).

Community engagement activities are currently used to help ensure ethical practice, build trust between parties, understand community perceptions of benefits and concerns regarding the planned research, develop effective and acceptable recruitment and follow-up strategies, review study materials and communicate results (242, 290-292). The definition for, and choice of, a community for a research project needs debate within and beyond the research team, and may in fact establish itself as a result of a project (293). This appeared to happen with the South African participants in the mixed-method study reported in Chapter 5, exhibited as a 'trial citizenship', and may in itself have influenced the measurement of some trial end points. So it will be important to take this into account when deciding how members of a community are to be involved in the topic. Including the public in determining the types of outcomes to be measured, and how, is low and dominated by North American and European groups (59). However, where the public/community have been included in these regions the results have been mutually beneficial (121, 294). This mirrors the gains from including patients as active partners in spontaneous ADR reporting post-registration, in that reports are not influenced by the prescriber's interpretation, and provide information about the effect of the ADR on the patient and their support network (66). It is imperative therefore to consider this for research conducted in low resource settings. As such, community members could be consulted about meaningful outcomes from the potential participants' perspectives, such as how they perceive certain side-effects, side-effects of concern within the local population, their potential burden on individuals and concerns about the impact of particular drugs on co-morbid conditions, concomitant medications or pregnancy (if relevant) (275, 295-299). They could also be involved in the co-development of simple messages to clarify for participants the rationale and importance of the information requested, to guide them as to the type(s) of data being asked about, and to highlight that

answers to questions are as important as study tests and examinations. It is important to also understand how to surmount more distal obstacles, such as the knowledge that one must not report traditional medicines as this is not allowed within the context of where the research is positioned. In-depth interviews with key informants may uncover innovative ways to negotiate this issue between patients and medical authorities.

Any questioning methods need to be feasible to use for both staff and participants. Feasibility will need to be assessed through pilot testing and by asking participants and staff about practicality and acceptability. Using online tools may not be workable in developing countries with low literacy and access to equipment. However, pictures, diaries and mobile phones may be options as there is precedence for their use in this region (74, 278, 300). Extensive questioning is unlikely to be acceptable for malaria treatment trials or in pregnancy registries because of the high turnover of patients at healthcare facilities. Using a question filter approach may make a tool more efficient (193), while linking questions about health with those about medicines should be explored as they are intuitive ways of questioning (e.g. if someone reports a headache - asking if they took something for that). It should also be acknowledged that, while researchers are permitted to limit the collection of ADRs later in the drug development pipeline under certain conditions, less severe but nonetheless important ADRs, or ADRs that primarily occur in vulnerable understudied populations, may yet have been detected. Malaria researchers should engage with others that are seeking to find ways to better represent the participants' adverse experiences of health in clinical trials and other studies. This would enable these issues to be raised and debated with regulatory authorities such that more investigators are aware of the challenges and successful approaches are shared across different therapeutic areas.

7.3 Conclusions

This thesis provides a much-needed African contribution to the literature concerned with how we obtain data required for the assessment of drug safety from participants in clinical trials and pharmacoepidemiology studies. Specifically, it has given a voice to participants, has extended the scope of the research beyond AEs to past medical histories and the use of medicines, and has considered pregnant women. It is clear that, when different methods are used to question participants about their health and use of medicines, the data detected will not be comparable. While some barriers to underreporting may be overcome

by a more comprehensive questioning method, others, such as worries about the adverse consequences of reporting, may not. This is because they reflect deep understandings or concerns about a participant's role in the research or their experiences within the wider healthcare environment. These issues may be more likely to be context-specific, whether at a personal, trial or structural level. Researchers should engage with their participant communities more about meaningful outcomes and mediate between them and relevant stakeholders to explore ways to overcome institutional or other barriers to reporting. The work has sensitised the malaria clinical research community about this under-studied issue and paved the way for further discussions and methodological work about optimal elicitation methods. It is critical that such work negotiates the tension between the positivist pharmacoepidemiological world and other perspectives on health. Thus, it is insufficient to compare participant reports as isolated fields in case record forms captured by different approaches and expect that there is an ideal tool that can be standardised for different contexts. Social science methods must be used to investigate the stories behind the data points, whether from the participants' or staff perspectives. Taking such an approach to this topic may result in the development of a framework for researchers to use when planning globally-relevant yet context-specific participant-reported data collection strategies. While this means that a method may not be completely standardised across all research sites, there would be a core of commonality to strengthen interpretation, comparisons and pooling of data from multiple studies.

REFERENCES

1. Guideline for Good Clinical Practice E6 (R1). International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use (ICH) 1996.
2. South African National Health Act. Available from: http://www.gov.za/sites/www.gov.za/files/a61-03.pdf. [Accessed 30th November 2015].
3. Dalkey N, Helmer, O. An experimental application of the Delphi method to the use of experts. Management Science. 1963;9:458-67.
4. Management of safety information from clinical trials. Geneva: Council for International Organizations of Medical Sciences (CIOMS) Working Group IV, 2005.
5. Ioannidis JP, Evans SJ, Gotzsche PC, O'Neill RT, Altman DG, Schulz K, et al. Better reporting of harms in randomized trials: an extension of the CONSORT statement. Annuals of Internernal Medicine. 2004;141(10):781-8.
6. Strom BL. Pharmacoepidemiology. 4th ed. Chichester, England: John Wiley & Sons; 2005.
7. Guidance for Industry: patient-reported outcome measures: use in medical product development to support labeling claims. US Food and Drug Administration, 2009.
8. The importance of pharmacovigilance: safety reporting of medicinal products. Geneva: World Health Organization, 2002.
9. Grzeskowiak LE, Gilbert AL, Morrison JL. Investigating outcomes associated with medication use during pregnancy: a review of methodological challenges and observational study designs. Reproductive Toxicology (Elmsford, NY). 2012;33(3):280-9.
10. Country and lending groups. World Bank; Available from: http://data.worldbank.org/about/country-and-lending-groups. [Accessed 30th November 2015].
11. Statistical principles for clinical trials E9. International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use (ICH) 1998.
12. Craven BM SG, Khan M. Chan TYK. Monitoring the safety of herbal medicines. Drug Safety. 1997;17:209-15.
13. Meyboom RH, Lindquist M, Flygare AK, Biriell C, Edwards IR. The value of reporting therapeutic ineffectiveness as an adverse drug reaction. Drug Safety. 2000;23(2):95-9.
14. Mehta U MJ, Duclos P, and Folb PI. Developing a national system for dealing with adverse events following immunization. Geneva: World Health Organization, 2000.
15. Mehta U, Allen E, Barnes KI. Establishing pharmacovigilance programs in resource-limited settings: the example of treating malaria. Expert Review of Clinical Pharmacology. 2010;3(4):509-25.
16. Mt-Isa S, Hallgreen CE, Wang N, Callreus T, Genov G, Hirsch I, et al. Balancing benefit and risk of medicines: a systematic review and classification of available methodologies. Pharmacoepidemiology and Drug Safety. 2014;23(7):667-78.
17. Posner J. Exploratory development. In: Griffin JP PJ, Barker GR, editor. The textbook of Pharmaceutical Medicine. 7 ed. London: John Wiley & Sons Ltd; 2013.
18. The future of drug safety: promoting and protecting the health of the public. Washington, DC: : Institute of Medicine (IOM), National Academy of Sciences, 2006.
19. Guidance for Industry: good pharmacovigilance practices and pharmacoepidemiologic assessment US Food and Drug Administration, 2005.
20. Barnes KI, Watkins WM, White NJ. Antimalarial dosing regimens and drug resistance. Trends in Parasitology. 2008;24(3):127-34.

21. Gomes MF, Faiz MA, Gyapong JO, Warsame M, Agbenyega T, Babiker A, et al. Pre-referral rectal artesunate to prevent death and disability in severe malaria: a placebo-controlled trial. Lancet. 2009;373(9663):557-66.
22. Ardal C, Rottingen JA. An open source business model for malaria. PloS One. 2015;10(2):e0117150.
23. Ndomondo-Sigonda M, Ambali A. The African Medicines Regulatory Harmonization Initiative: rationale and benefits. Clinical Pharmacology and Therapeutics. 2011;89(2):176-8.
24. Drug development research in resource-limited countries. Geneva: CIOMS/WHO Working Group, 2005.
25. Hodel EM, Kabanywanyi AM, Malila A, Zanolari B, Mercier T, Beck HP, et al. Residual antimalarials in malaria patients from Tanzania--implications on drug efficacy assessment and spread of parasite resistance. PloS One. 2009;4(12):e8184.
26. Skalli S, Zaid A, Soulaymani R. Drug interactions with herbal medicines. Therapeutic Drug Monitoring. 2007;29(6):679-86.
27. Lang T, Hughes D, Kanyok T, Kengeya-Kayondo J, Marsh V, Haaland A, et al. Beyond registration--measuring the public-health potential of new treatments for malaria in Africa. The Lancet Infectious diseases. 2006;6(1):46-52.
28. World Malaria Report Geneva: World Health Organization, 2014.
29. Guidelines for the Treatment of Malaria. Geneva: World Health Organization, 2015.
30. Ramos-Martin V, Gonzalez-Martinez C, Mackenzie I, Schmutzhard J, Pace C, Lalloo DG, et al. Neuroauditory toxicity of artemisinin combination therapies-have safety concerns been addressed? The American Journal of Tropical Medicine and Hygiene. 2014;91(1):62-73.
31. Manyando C, Njunju EM, Virtanen M, Hamed K, Gomes M, Van Geertruyden JP. Exposure to artemether-lumefantrine (Coartem) in first trimester pregnancy in an observational study in Zambia. Malaria Journal. 2015;14:77.
32. Tanner M, Greenwood B, Whitty CJ, Ansah EK, Price RN, Dondorp AM, et al. Malaria eradication and elimination: views on how to translate a vision into reality. BMC Medicine. 2015;13:167.
33. Dondorp AM, Nosten F, Yi P, Das D, Phyo AP, Tarning J, et al. Artemisinin resistance in Plasmodium falciparum malaria. The New England Journal of Medicine. 2009;361(5):455-67.
34. Baiden R, Oduro A, Halidou T, Gyapong M, Sie A, Macete E, et al. Prospective observational study to evaluate the clinical safety of the fixed-dose artemisinin-based combination Eurartesim(R) (dihydroartemisinin/piperaquine), in public health facilities in Burkina Faso, Mozambique, Ghana, and Tanzania. Malaria Journal. 2015;14:160.
35. Mosha D, Mazuguni F, Mrema S, Sevene E, Abdulla S, Genton B. Safety of artemether-lumefantrine exposure in first trimester of pregnancy: an observational cohort. Malaria Journal. 2014;13:197.
36. Dodoo AN, Fogg C, Nartey ET, Ferreira GL, Adjei GO, Kudzi W, et al. Profile of adverse events in patients receiving treatment for malaria in urban Ghana: a cohort-event monitoring study. Drug Safety. 2014;37(6):433-48.
37. ACT Consortium. Available from: http://www.actconsortium.org/. [Accessed 15th August 2015].
38. Posner J WS. Objectives and design of clinical trials. In: Griffin JP PJ, Barker GR, editor. The Textbook of Pharmaceutical Medicine. 7 ed. London: John Wiley & Sons Ltd; 2013.
39. General considerations for clinical trials E8. International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use (ICH) 1997.

40. Schaefer C, Ornoy A, Clementi M, Meister R, Weber-Schoendorfer C. Using observational cohort data for studying drug effects on pregnancy outcome--methodological considerations. Reproductive Toxicology (Elmsford, NY). 2008;26(1):36-41.
41. Howard TB, Tassinari MS, Feibus KB, Mathis LL. Monitoring for teratogenic signals: pregnancy registries and surveillance methods. American Journal of Medical Genetics Part C, Seminars in Medical Genetics. 2011;157C(3):209-14.
42. Kennedy DL, Uhl K, Kweder SL. Pregnancy exposure registries. Drug Safety. 2004;27(4):215-28.
43. Mehta U, Clerk C, Allen E, Yore M, Sevene E, Singlovic J, et al. Protocol for a drugs exposure pregnancy registry for implementation in resource-limited settings. BMC Pregnancy and Childbirth. 2012;12:89.
44. Pharmacovigilance planning E2E International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use (ICH) 2005.
45. Harmark L, van Grootheest AC. Pharmacovigilance: methods, recent developments and future perspectives. European Journal of Clinical Pharmacology. 2008;64(8):743-52.
46. Horvath CJ, Milton MN. The TeGenero incident and the Duff Report conclusions: a series of unfortunate events or an avoidable event? Toxicologic Pathology. 2009;37(3):372-83.
47. Guidelines for phase 1 clinical trials. London: Association of the British Pharmaceutical Industry (ABPI), 2012.
48. Edgerly M, Fojo T. Is there room for improvement in adverse event reporting in the era of targeted therapies? Journal of the National Cancer Institute. 2008;100(4):240-2.
49. Martinez E, editor. Safety and tolerability of currently available antiretroviral agents. 2nd International AIDS Society Conference on HIV Pathogenesis, Treatment and Prevention; 2003.
50. Lewiecki EM. Risk Communication and Shared Decision Making in the Care of Patients With Osteoporosis. Journal of Clinical Densitometry. 2010;13(4):335-45.
51. Winkler EC, Hiddemann W, Marckmann G. Evaluating a patient's request for life-prolonging treatment: an ethical framework. Journal of Medical Ethics. 2012;38(11):647-51.
52. Singh S, Loke YK. Drug safety assessment in clinical trials: methodological challenges and opportunities. Trials. 2012;13:138.
53. Njomo DW, Amuyunzu-Nyamongo M, Magambo JK, Njenga SM. The role of personal opinions and experiences in compliance with mass drug administration for lymphatic filariasis elimination in Kenya. PloS One. 2012;7(11):e48395.
54. Chen I, Poirot E, Newman M, Kandula D, Shah R, Hwang J, et al. An assessment of the supply, programmatic use, and regulatory issues of single low-dose primaquine as a Plasmodium falciparum gametocytocide for sub-Saharan Africa. Malaria Journal. 2015;14:204.
55. Structure and content of clinical study reports E3. International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use (ICH) 1996.
56. Davis C, Ralevski E, Kennedy SH, Neitzert C. The role of personality factors in the reporting of side effect complaints to moclobemide and placebo: a study of healthy male and female volunteers. Journal of Clinical Psychopharmacology. 1995;15(5):347-52.
57. Lin KM, Smith MW, Ortiz V. Culture and psychopharmacology. The Psychiatric Cinics of North America. 2001;24(3):523-38.
58. Etkin NL. "Side effects": cultural constructions and reinterpretations of Western pharmaceuticals. Medical Anthropology. 1992;6:99-113.

59. Gargon E, Gurung B, Medley N, Altman DG, Blazeby JM, Clarke M, et al. Choosing important health outcomes for comparative effectiveness research: a systematic review. PloS One. 2014;9(6):e99111.
60. Loke YK, Price D, Herxheimer A, Cochrane Adverse Effects Methods G. Systematic reviews of adverse effects: framework for a structured approach. BMC Medical Research Methodology. 2007;7:32.
61. Reviewer guide: conducting a clinical safety review of a new product application and preparing a report in the review. US Food and Drug Administration, 2005.
62. Ting Lee M-L LR. Meta-analysis of drug safety data with logistic regression. Drug Information Journal. 1997;31:1189-93.
63. Pope A, Adams C, Paton C, Weaver T, Barnes TR. Assessment of adverse effects in clinical studies of antipsychotic medication: survey of methods used. The British Journal of Psychiatry: the Journal of Mental Science. 2010;197(1):67-72.
64. Barbour V, Clark J, Jones S, Norton M, Simpson P, Veitch E. Why drug safety should not take a back seat to efficacy. PLoS Medicine. 2011;8(9):e1001097.
65. Ioannidis JP. Adverse events in randomized trials: neglected, restricted, distorted, and silenced. Archives of Internal Medicine. 2009;169(19):1737-9.
66. Herxheimer A. Pharmacovigilance on the turn? Adverse reactions methods in 2012. The International Journal of Risk & Safety in Medicine. 2013;25(1):29-31.
67. Doda Z. Introduction to Sociocultural Anthropology. Ethiopia: Debub University; 2005. Available from: http://www.cartercenter.org/resources/pdfs/health/ephti/library/lecture_notes/health_science_students/ln_socio_anthro_final.pdf. [Accessed 30th November 2015].
68. Kleinman A, Benson P. Anthropology in the clinic: the problem of cultural competency and how to fix it. PLoS Medicine. 2006;3(10):e294.
69. Takeuchi T, Nakao M, Kohno M, Hatano M, Niimi M, Yano E. Development of a questionnaire to assess 'Hie' symptoms using an evidence-based analysis. Environmental Health and Preventive Medicine. 2008;13(6):338-44.
70. Faasse K, Grey A, Horne R, Petrie KJ. High perceived sensitivity to medicines is associated with higher medical care utilisation, increased symptom reporting and greater information-seeking about medication. Pharmacoepidemiology and Drug Safety. 2015;24(6):592-9.
71. Bjornsdottir I, Almarsdottir AB, Traulsen JM. The lay public's explicit and implicit definitions of drugs. Research in Social & Administrative Pharmacy. 2009;5(1):40-50.
72. Armstrong D, Lilford R, Ogden J, Wessely S. Health-related quality of life and the transformation of symptoms. Sociology of Health & Illness. 2007;29(4):570-83.
73. Haaland A, Molyneux S, Marsh V. Quality information in field research: Training manual on practical communication skills for field researchers and project personnel. 2006;World Health Organization on behalf of the Special Programme for Research and Training in Tropical Diseases.
74. Davies EC, Chandler CI, Innocent SH, Kalumuna C, Terlouw DJ, Lalloo DG, et al. Designing adverse event forms for real-world reporting: participatory research in Uganda. PloS One. 2012;7(3):e32704.
75. Haaland A. Reporting with pictures: a concept paper for researchers and health policy decision-makers. 2000;World Health Organization on behalf of the Special Programme for Research and Training in Tropical Diseases.
76. Woodsong C, Mutsambi JM, Ntshele S, Modikoe P. Community and research staff collaboration for development of materials to inform microbicide study participants in Africa. Journal of the International AIDS Society. 2014;17(3Suppl 2):19156.
77. Adams V, Miller S, Craig S, Samen A, Nyima, Sonam, et al. The challenge of cross-cultural clinical trials research: case report from the Tibetan Autonomous Region, People's Republic of China. Medical Anthropology Quarterly. 2005;19(3):267-89.

78. Mhlongo SW, Mbokazi AJ. Ethnic differences in risks of adverse reactions: biomedical approach is insufficient to explain ethnic differences. British Medical Journal (Clinical Research ed). 2006;332(7554):1393.
79. Korfage IJ, Hak T, de Koning HJ, Essink-Bot ML. Patients' perceptions of the side-effects of prostate cancer treatment--a qualitative interview study. Social Science & Medicine (1982). 2006;63(4):911-9.
80. Lee S. Side effects of chronic lithium therapy in Hong Kong Chinese: an ethnopsychiatric perspective. Cultural Medical Psychiatry. 1993;17(3):301-20.
81. Rief W, Nestoriuc Y, von Lilienfeld-Toal A, Dogan I, Schreiber F, Hofmann SG, et al. Differences in adverse effect reporting in placebo groups in SSRI and tricyclic antidepressant trials: a systematic review and meta-analysis. Drug Safety. 2009;32(11):1041-56.
82. Myers MG, Cairns JA, Singer J. The consent form as a possible cause of side effects. Clinical Pharmacology and Therapeutics. 1987;42(3):250-3.
83. Basch E. The missing voice of patients in drug-safety reporting. The New England Journal of Medicine. 2010;362(10):865-9.
84. The International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use (ICH). Available from: http://www.ich.org/home.html. [Accessed 15th August 2015].
85. Guidance for Industry: premarketing risk assessment. US Food and Drug Administration, 2005.
86. Draft guidance for Industry: suicidal ideation and behavior: prospective assessment of occurrence in clinical trials US Food and Drug Administration, 2012.
87. Guideline on clinical investigation of medicinal products in the treatment of depression. London: European Medicines Agency, 2013.
88. Draft guidance for Industry: evaluating drug effects on the ability to operate a motor vehicle US Food and Drug Administration, 2015.
89. Draft guidance for Industry: assessment of abuse potential of drugs. US Food and Drug Administration, 2010.
90. Reflection paper on the use of patient reported outcome (PRO) measures in oncology studies. London: European Medicines Agency Oncology Working Party, 2014.
91. Draft guidance for Industry: determining the extent of safety data collection needed in late stage premarket and postapproval clinical investigations US Food and Drug Administration, 2012.
92. Staedke SG, Jagannathan P, Yeka A, Bukirwa H, Banek K, Maiteki-Sebuguzi C, et al. Monitoring antimalarial safety and tolerability in clinical trials: a case study from Uganda. Malaria Journal. 2008;7:107.
93. Clarke M. Standardising outcomes for clinical trials and systematic reviews. Trials. 2007;8:39.
94. Williamson PR, Altman DG, Blazeby JM, Clarke M, Devane D, Gargon E, et al. Developing core outcome sets for clinical trials: issues to consider. Trials. 2012;13:132.
95. Sinha IP, Smyth RL, Williamson PR. Using the Delphi technique to determine which outcomes to measure in clinical trials: recommendations for the future based on a systematic review of existing studies. PLoS Medicine. 2011;8(1):e1000393.
96. Methods for surveillance of antimalarial drug efficacy. Geneva: World Health Organization, 2009.
97. Guidelines for good pharmacoepidemiology practice. International Society for Pharmacoepidemiology, 2007.
98. Reviewer guidance: evaluating the risks of drug exposure in human pregnancies US Food and Drug Administration, 2005.
99. Guidance for Industry: establishing pregnancy exposure registries. US Food and Drug Administration, 2002.

100. Friedman LM FC, DeMets D. Assessing and reporting adverse events. Fundamentals of Clinical Trials. New York: Springer-Verlag 2010. p. 215-31.
101. Loke YK, Derry S. Reporting of adverse drug reactions in randomised controlled trials - a systematic survey. BMC Clinical Pharmacology. 2001;1:3.
102. Hodkinson A, Kirkham JJ, Tudur-Smith C, Gamble C. Reporting of harms data in RCTs: a systematic review of empirical assessments against the CONSORT harms extension. British Medical Journal Open. 2013;3(9):e003436.
103. Jordan S, Knight J, Pointon D. Monitoring adverse drug reactions: scales, profiles, and checklists. International Nursing Review. 2004;51(4):208-21.
104. Barnes TR. The Barnes Akathisia Rating Scale--revisited. Journal of Psychopharmacology (Oxford, England). 2003;17(4):365-70.
105. Levine J, Schooler NR. SAFTEE: a technique for the systematic assessment of side effects in clinical trials. Psychopharmacology Bulletin. 1986;22(2):343-81.
106. Rabkin JG, Markowitz JS, Ocepek-Welikson K, Wager SS. General versus systematic inquiry about emergent clinical events with SAFTEE: implications for clinical research. Journal of Clinical Psychopharmacology. 1992;12(1):3-10.
107. Day JC, Wood G, Dewey M, Bentall RP. A self-rating scale for measuring neuroleptic side-effects. Validation in a group of schizophrenic patients. The British Journal of Psychiatry: the Journal of Mental Science. 1995;166(5):650-3.
108. Ohlsen RI, Williamson R, Yusufi B, Mullan J, Irving D, Mukherjee S, et al. Interrater reliability of the Antipsychotic Non-Neurological Side-Effects Rating Scale measured in patients treated with clozapine. Journal of Psychopharmacology. 2008;22(3):323-9.
109. Byerly MJ, Nakonezny PA, Fisher R, Magouirk B, Rush AJ. An empirical evaluation of the Arizona sexual experience scale and a simple one-item screening test for assessing antipsychotic-related sexual dysfunction in outpatients with schizophrenia and schizoaffective disorder. Schizophrenia Research. 2006;81(2-3):311-6.
110. Lindstrom E, Lewander T, Malm U, Malt UF, Lublin H, Ahlfors UG. Patient-rated versus clinician-rated side effects of drug treatment in schizophrenia. Clinical validation of a self-rating version of the UKU Side Effect Rating Scale (UKU-SERS-Pat). Nordic Journal of Psychiatry. 2001;55 Suppl 44:5-69.
111. Ko YH, Na KS, Kim CE, Kim SH, Jeon YW, Yi JS, et al. The effectiveness of cross-tapering switching to ziprasidone in patients with schizophrenia or schizoaffective disorder. Psychiatry Investigation. 2014;11(4):459-66.
112. Greenhill LL, Vitiello B, Fisher P, Levine J, Davies M, Abikoff H, et al. Comparison of increasingly detailed elicitation methods for the assessment of adverse events in pediatric psychopharmacology. Journal of the American Academy of Child and Adolescent Psychiatry. 2004;43(12):1488-96.
113. Waddell L, Taylor M. A new self-rating scale for detecting atypical or second-generation antipsychotic side effects. Journal of Psychopharmacology. 2008;22(3):238-43.
114. Nugent KL, Spahr E, Toroney J, Mojtabai R, Nettles C, Turner LW, et al. Interrater reliability of the modified monitoring of side effects scale for assessment of adverse effects of psychiatric medication in clinical and research settings. Journal of Clinical Psychopharmacology. 2015;35(3):324-8.
115. Rosenberg KP, Bleiberg KL, Koscis J, Gross C. A survey of sexual side effects among severely mentally ill patients taking psychotropic medications: impact on compliance. Journal of Sex & Marital Therapy. 2003;29(4):289-96.
116. Common Terminology Criteria for Adverse Events (CTCAE) v.4. National Cancer Institute. [Accessed 15th August 2015].
117. Atkinson TM, Li Y, Coffey CW, Sit L, Shaw M, Lavene D, et al. Reliability of adverse symptom event reporting by clinicians. Quality of Life Research: an international journal of quality of life aspects of treatment, care and rehabilitation. 2012;21(7):1159-64.

118. Pakhomov SV, Jacobsen SJ, Chute CG, Roger VL. Agreement between patient-reported symptoms and their documentation in the medical record. The American Journal of Managed Care. 2008;14(8):530-9.
119. Basch E, Reeve BB, Mitchell SA, Clauser SB, Minasian LM, Dueck AC, et al. Development of the National Cancer Institute's patient-reported outcomes version of the common terminology criteria for adverse events (PRO-CTCAE). Journal of the National Cancer Institute. 2014;106(9).
120. Tugwell P, Boers M, Brooks P, Simon L, Strand V, Idzerda L. OMERACT: an international initiative to improve outcome measurement in rheumatology. Trials. 2007;8:38.
121. Kirwan JR, Minnock P, Adebajo A, Bresnihan B, Choy E, de Wit M, et al. Patient perspective: fatigue as a recommended patient centered outcome measure in rheumatoid arthritis. The Journal of Rheumatology. 2007;34(5):1174-7.
122. Woodworth T, Furst DE, Alten R, Bingham CO, 3rd, Yocum D, Sloan V, et al. Standardizing assessment and reporting of adverse effects in rheumatology clinical trials II: the Rheumatology Common Toxicity Criteria v.2.0. The Journal of Rheumatology. 2007;34(6):1401-14.
123. Lassere MN, Johnson KR, Van Santen S, Carlton K, Rappo J, Michael R, et al. Generic patient self-report and investigator report instruments of therapeutic safety and tolerability. The Journal of Rheumatology. 2005;32(10):2033-6.
124. Carr A. Improvement of the study, analysis, and reporting of adverse events associated with antiretroviral therapy. Lancet. 2002;360(9326):81-5.
125. Chowers MY, Gottesman BS, Leibovici L, Pielmeier U, Andreassen S, Paul M. Reporting of adverse events in randomized controlled trials of highly active antiretroviral therapy: systematic review. The Journal of Antimicrobial Chemotherapy. 2009;64(2):239-50.
126. Chrysant SG. Proactive compared with passive adverse event recognition: calcium channel blocker-associated edema. Journal of Clinical Hypertension (Greenwich, Conn). 2008;10(9):716-22.
127. Cornelius VR, Sauzet O, Williams JE, Ayis S, Farquhar-Smith P, Ross JR, et al. Adverse event reporting in randomised controlled trials of neuropathic pain: considerations for future practice. Pain. 2013;154(2):213-20.
128. Edwards JE, McQuay HJ, Moore RA, Collins SL. Reporting of adverse effects in clinical trials should be improved: lessons from acute postoperative pain. Journal of Pain and Symptom Management. 1999;18(6):427-37.
129. Foster JM, van der Molen T, Caeser M, Hannaford P. The use of questionnaires for measuring patient-reported side effects of drugs: its importance and methodological challenges. Pharmacoepidemiol Drug Safety. 2008;17(3):278-96.
130. Greenhill LL, Vitiello B, Riddle MA, Fisher P, Shockey E, March JS, et al. Review of safety assessment methods used in pediatric psychopharmacology. Journal of the American Academy of Child and Adolescent Psychiatry. 2003;42(6):627-33.
131. Calis KA, Daniels CE. Safety reporting in clinical trials. JAMA. 2001;285(16):2076; author reply 7-8.
132. Hughes S, Cohen D. Understanding the assessment of psychotropic drug harms in clinical trials to improve social workers' role in medication monitoring. Social Work. 2010;55(2):105-15.
133. Katz NP. The measurement of symptoms and side effects in clinical trials of chronic pain. Contemporary Clinical Trials. 2012;33(5):903-11.
134. Vida S, Looper K. Precision and comparability of adverse event rates of newer antidepressants. Journal of Clinical Psychopharmacology. 1999;19(5):416-26.
135. Rief W GJ, Barsky AJ Generic assessment of side effects. 2009. Available from: http://gase-scale.com/Instruction_GASE.pdf. [Accessed 15th August 2015].

136. Streptomycin treatment of pulmonary tuberculosis. British Medical Journal. 1948;2(4582):769-82.
137. Glaser EM, Whittow GC. Experimental errors in clinical trials. Clinical Science (London, England: 1979). 1954;13(2):199-210.
138. Avery CW, Ibelle BP, Allison B, Mandell N. Systematic errors in the evaluation of side effects. American Journal of Psychiatry. 1967;123(7):875-8.
139. Bent S, Padula A, Avins AL. Brief communication: better ways to question patients about adverse medical events: a randomized, controlled trial. Annals of Internal Medicine. 2006;144(4):257-61.
140. Allen EN, Mushi AK, Massawe IS, Vestergaard LS, Lemnge M, Staedke SG, et al. How experiences become data: the process of eliciting adverse event, medical history and concomitant medication reports in antimalarial and antiretroviral interaction trials. BMC Medical Research Methodology. 2013;13:140.
141. Barber BL, Santanello NC. Relating spontaneous adverse experience reports to scores on a questionnaire querying tolerability. International Journal of Clinical Pharmacology and Therapeutics. 1995;33(11):598-604.
142. Barrowman JA, Herxheimer A, Kits TP. Unwanted effects of pentagastrin. Clinical Pharmacology and Therapeutics. 1970;11(6):862-8.
143. Borghi C, Pallavini G, Comi D, Grillo G, Lombardo M, Mantero O, et al. Comparison of three different methods of monitoring unwanted effects during antihypertensive therapy. International Journal of Clinical Pharmacology, Therapy, and Toxicology. 1984;22(6):324-8.
144. Brent DA, Emslie GJ, Clarke GN, Asarnow J, Spirito A, Ritz L, et al. Predictors of spontaneous and systematically assessed suicidal adverse events in the treatment of SSRI-resistant depression in adolescents (TORDIA) study. The American Journal of Psychiatry. 2009;166(4):418-26.
145. Ciccolunghi SN, Chaudri HA. A methodological study of some factors influencing the reporting of symptoms. Journal of Clinical Pharmacology. 1975;15(7):496-505.
146. Downing RW, Rickels K, Meyers F. Side reactions in neurotics. I. A comparison of two methods of assessment. The Journal of Clinical Pharmacology and the Journal of New Drugs. 1970;10(5):289-97.
147. Huskisson EC, Wojtulewski JA. Measurement of side effects of drugs. British Medical Journal. 1974;2(5921):698-9.
148. Jacobson AF, Goldstein BJ, Dominguez RA, Steinbook RM. Interrater agreement and reliability measures of SAFTEE: general inquiry vs. systematic inquiry. Psychopharmacology Bulletin. 1987;23(1):97-101.
149. Kruft B, Nelson LA, Stewart JA, Stewart WC. Adverse event reporting. Ophthalmology. 2007;114(7):1420.
150. Landén M, Högberg P, Thase ME. Incidence of sexual side effects in refractory depression during treatment with citalopram or paroxetine. The Journal of Clinical Psychiatry. 2005; 66(1):100-6.
151. Lundberg PK. Assessment of drugs' side effects: Visual Analogue Scale versus check-list format. Perceptual and Motor Skills. 1980;50(3 Pt 2):1067-73.
152. Monteiro WO, Noshirvani HF, Marks IM, Lelliott PT. Anorgasmia from clomipramine in obsessive-compulsive disorder. A controlled trial. The British Journal of Psychiatry: the Journal of Mental Science. 1987;151:107-12.
153. Nicholls DP, Husaini MH, Bulpitt CJ, Stephens MD, Butler AG. Comparison of labetalol and propranolol in hypertension. British Journal of Clinical Pharmacology. 1980;9(3):233-7.
154. O'Connell K, Davis AR, Kerns J. Oral contraceptives: side effects and depression in adolescent girls. Contraception. 2007;75(4):299-304.

155. Os I, Bratland B, Dahlöf B, Gisholt K, Syvertsen JO, Tretli S. Female preponderance for lisinopril-induced cough in hypertension. American Journal of Hypertension [Internet]. 1994; 7(11):[1012-5].
156. Perez-Lloret S, Rey MV, Fabre N, Ory F, Spampinato U, Montastruc JL, et al. Do Parkinson's disease patients disclose their adverse events spontaneously? European Journal of Clinical Pharmacology. 2012;68(5):857-65.
157. Reilly MC, Zbrozek AS. Assessing the responsiveness of a quality-of-life instrument and the measurement of symptom severity in essential hypertension. PharmacoEconomics. 1992;2(1):54-66.
158. Rosenthal J, Bahrmann H, Benkert K, Baumgart P, Bonner G, Klein G, et al. Analysis of adverse effects among patients with essential hypertension receiving an ACE inhibitor or a beta-blocker. Cardiology. 1996;87(5):409-14.
159. Sheftell FD, Feleppa M, Tepper SJ, Rapoport AM, Ciannella L, Bigal ME. Assessment of adverse events associated with triptans--methods of assessment influence the results. Headache. 2004;44(10):978-82.
160. Spilker A KJ. Comparison of symptoms elicited by checklist and fill-in-the-blank questionnaires. Pharmacoepidemiology News 1987;22:272-7.
161. Török E, Kósa E, Szontagh M, Juvancz P, Eggenhofer J, Borvendég J. Comparison of methods to detect side-effect on clinical application of chloranolol, a beta-adrenergic receptor inhibitor. European Journal of Clinical Pharmacology [Internet]. 1984; 26(5):[555-62].
162. Wallander MA, Dimenas E, Svardsudd K, Wiklund I. Evaluation of three methods of symptom reporting in a clinical trial of felodipine. European Journal of Clinical Pharmacology. 1991;41(3):187-96.
163. Wallin J, Sjovall J. Detection of adverse drug reactions in a clinical trial using two types of questioning. Clinical Therapeutics. 1981;3(6):450-2.
164. Yeo WW, Maclean D, Richardson PJ, Ramsay LE. Cough and enalapril: assessment by spontaneous reporting and visual analogue scale under double-blind conditions. British Journal of Clinical Charmacology [Internet]. 1991; 31(3):[356-9 pp.].
165. Wernicke JF, Faries D, Milton D, Weyrauch K. Detecting treatment emergent adverse events in clinical trials: a comparison of spontaneously reported and solicited collection methods. Drug Safety. 2005;28(11):1057-63.
166. Herxheimer A. Letter: measurement of side effects of drugs. British Medical Journal. 1974;3(5928):470.
167. Ioannidis JP, Mulrow CD, Goodman SN. Adverse events: the more you search, the more you find. Annals of Internal Medicine. 2006;144(4):298-300.
168. Banerjee A, Okun S, Edwards IR, Wicks P, Smith M, Mayall S, et al. Patient-Reported Outcome Measures in Safety Event Reporting: PROSPER Consortium Guidance. Drug Safety. 2013;36(12):1129-49.
169. List of pregnancy exposure registries. US Food and Drug Administration; Available from: http://www.fda.gov/ScienceResearch/SpecialTopics/WomensHealthResearch/ucm134848.htm. [Accessed 15th August 2015].
170. Chambers CD, Braddock SR, Briggs GG, Einarson A, Johnson YR, Miller RK, et al. Postmarketing surveillance for human teratogenicity: a model approach. Teratology. 2001;64(5):252-61.
171. Mosha D, Mazuguni F, Mrema S, Abdulla S, Genton B. Medication exposure during pregnancy: a pilot pharmacovigilance system using health and demographic surveillance platform. BMC Pregnancy and Childbirth. 2014;14:322.
172. Mitchell AA, Cottler LB, Shapiro S. Effect of questionnaire design on recall of drug exposure in pregnancy. American Journal of Epidemiology. 1986;123(4):670-6.

173. West SL SB, Poll C. Validity of pharmacoepidemiologic drug and diagnosis data. In: BL S, editor. Pharmacoepidemiology. 4 ed. Chichester, England: John Wiley & Sons; 2005. p. 709-65.
174. Gama H, Correia S, Lunet N. Questionnaire design and the recall of pharmacological treatments: a systematic review. Pharmacoepidemiology and Drug Safety. 2009;18(3):175-87.
175. Karim QA, Kharsany AB, Naidoo K, Yende N, Gengiah T, Omar Z, et al. Co-enrollment in multiple HIV prevention trials - experiences from the CAPRISA 004 Tenofovir gel trial. Contemporary Clinical Trials. 2011;32(3):333-8.
176. O'Leary Z. The essential guide to doing your research project. 2e ed: SAGE Publications Ltd. London, UK; 2014.
177. Conrad P, Barker KK. The social construction of illness: key insights and policy implications. Journal of Health and Social Behavior. 2010;51 Suppl:S67-79.
178. Merleau-Ponty M. From phenomenology of perception. In: Lock M, Farquhar J, editors. Beyond the body proper: reading the anthropology of material life. Durham, NC: Duke University Press; 2007. p. 133-49
179. Lock M, Farquhar J. Beyond the body proper: reading the anthropology of material life. Durham, NC: Duke University Press; 2007.
180. Pool R, Montgomery CM, Morar NS, Mweemba O, Ssali A, Gafos M, et al. A mixed methods and triangulation model for increasing the accuracy of adherence and sexual behaviour data: the Microbicides Development Programme. PloS One. 2010;5(7):e11600.
181. Assessment of the safety of artemisinin compounds in pregnancy: report of two joint informal consultations convened in 2006. France. France: World Health Organization, 2007.
182. Huang HY, Andrews E, Jones J, Skovron ML, Tilson H. Pitfalls in meta-analyses on adverse events reported from clinical trials. Pharmacoepidemiology and Drug Safety. 2011;20(10):1014-20.
183. Allen EN, Barnes KI, Mushi A, Massawe I, Staedke SG, Mehta U, et al. Eliciting harms data from trial participants: how perceptions of illness and treatment mediate recognition of relevant information to report. Trials. 2011;12.
184. Horsley T, Dingwall O, Sampson M. Checking reference lists to find additional studies for systematic reviews. The Cochrane Database of Systematic Reviews. 2011(8):MR000026.
185. Scherer RW, Langenberg P, von Elm E. Full publication of results initially presented in abstracts. The Cochrane Database of Systematic Reviews. 2007(2):MR000005.
186. Cochrane handbook for systematic reviews of interventions. 5.1.0 ed: The Cochrane Collaboration; 2011.
187. Young T, Hopewell S. Methods for obtaining unpublished data. The Cochrane Database of Systematic Reviews. 2011(11):MR000027.
188. Golder S, Loke YK, Bland M. Meta-analyses of adverse effects data derived from randomised controlled trials as compared to observational studies: methodological overview. PLoS Medicine. 2011;8(5):e1001026.
189. Higgins JP, Thompson SG. Quantifying heterogeneity in a meta-analysis. Statistics in Medicine. 2002;21(11):1539-58.
190. Sterne JA, Egger M. Funnel plots for detecting bias in meta-analysis: guidelines on choice of axis. Journal of Clinical Epidemiology. 2001;54(10):1046-55.
191. Popay J, Roberts, H., Sowden, A., Petticrew, M., Arai, L., Rodgers, M., Britten, N., Guidance on the Conduct of Narrative Synthesis in Systematic Reviews: A Product from the ESRC Methods Programme. 2006.
192. de Vries ST, Haaijer-Ruskamp FM, de Zeeuw D, Denig P. The validity of a patient-reported adverse drug event questionnaire using different recall periods. Quality of Life

Research: an international journal of quality of life aspects of treatment, care and rehabilitation. 2014;23(9):2439-45.
193. de Vries ST, Mol PG, de Zeeuw D, Haaijer-Ruskamp FM, Denig P. Development and Initial Validation of a Patient-Reported Adverse Drug Event Questionnaire. Drug Safety. 2013.
194. Hermans L, Deblander A, De Keyser P, Scheys I, Lesaffre E, Westelinck KJ. At equipotent doses, isradipine is better tolerated than amlodipine in patients with mild-to-moderate hypertension: A double-blind, randomized, parallel-group study. British Journal of Clinical Pharmacology. 1994;38(4):335-40.
195. Heaven B, Murtagh M, Rapley T, May C, Graham R, Kaner E, et al. Patients or research subjects? A qualitative study of participation in a randomised controlled trial of a complex intervention. Patient Education and Counseling. 2006;62(2):260-70.
196. Paterson C, Zheng Z, Xue C, Wang Y. "Playing their parts": the experiences of participants in a randomized sham-controlled acupuncture trial. Journal of Alternative and Complementary Medicine (New York, NY). 2008;14(2):199-208.
197. Scott C, Walker J, White P, Lewith G. Forging convictions: the effects of active participation in a clinical trial. Social Science & Medicine (1982). 2011;72(12):2041-8.
198. Price RN, Dorsey G, Ashley EA, Barnes KI, Baird JK, d'Alessandro U, et al. World Antimalarial Resistance Network I: clinical efficacy of antimalarial drugs. Malaria Journal. 2007;6:119.
199. Golder S, Loke YK. Failure or success of electronic search strategies to identify adverse effects data. Journal of the Medical Library Association : JMLA. 2012;100(2):130-4.
200. DMID:Toxicity tables. National Institute of Allergies and Infectious Diseases, 2007.
201. Kohl KS, Gidudu J, Bonhoeffer J, Braun MM, Buettcher M, Chen RT, et al. The development of standardized case definitions and guidelines for adverse events following immunization. Vaccine. 2007;25(31):5671-4.
202. Rief W, Barsky AJ, Glombiewski JA, Nestoriuc Y, Glaesmer H, Braehler E. Assessing general side effects in clinical trials: reference data from the general population. Pharmacoepidemiol Drug Safety. 2011;20(4):405-15.
203. Malaria in Pregnancy Consortium. Available from: http://mip-consortium.org/. [Accessed 15th August 2015].
204. PubMed. Available from: http://www.ncbi.nlm.nih.gov/pubmed/. Accessed 15th August 2015.
205. ClinicalTrials.gov. Available from: https://clinicaltrials.gov/. [Accessed 15th August 2015].
206. The Pan African Clinical Trials Registry. Available from: http://www.pactr.org/. [Accessed 15th August 2015].
207. Database of African health researchers. Initiative to Strengthen Health Research Capacity in Africa (ISHReca); Available from: http://ishreca.tropika.net/news/database-of-african-health-researchers. [Accessed 26th October 2013].
208. WorldWide Antimalarial Resistance Network. Available from: http://www.wwarn.org/. [Accessed 15th August 2015].
209. Global Health Trials. [cited 2015 15th August]; Available from: https://globalhealthtrials.tghn.org/. [Accessed 15th August 2015].
210. SurveyGizmo®. Available from: http://www.surveygizmo.com/. [Accessed 15th August 2015].
211. Strauss A CJ. Basics of qualitative research: grounded theory procedures and techniques. London: Sage; 1990.
212. Toxicity grading scale for determining the severity of adverse events. World Health Organization, 2003.

213. Saathoff E, Schneider P, Kleinfeldt V, Geis S, Haule D, Maboko L, et al. Laboratory reference values for healthy adults from southern Tanzania. Tropical Medicine & International Health. 2008;13(5):612-25.
214. WHO-UMC system for standardised case causality assessment. World Health Organisation-Uppsala Monitoring Centre.
215. Hill AB. The environment and disease: association or causation? Proceedings of the Royal Society of Medicine. 1965;58:295-300.
216. The Medical Dictionary for Regulatory Activities. Available from: http://meddramsso.com/. [Accessed 15th August 2015].
217. Sinclair D, Zani B, Donegan S, Olliaro P, Garner P. Artemisinin-based combination therapy for treating uncomplicated malaria. The Cochrane Database of Systematic Reviews. 2009(3):CD007483.
218. Haidich AB, Birtsou C, Dardavessis T, Tirodimos I, Arvanitidou M. The quality of safety reporting in trials is still suboptimal: survey of major general medical journals. Journal of Clinical Epidemiology. 2011;64(2):124-35.
219. Cahill P, Papageorgiou A. Triadic communication in the primary care paediatric consultation: a review of the literature. The British Journal of General Practice. 2007;57(544):904-11.
220. Trotti A, Colevas AD, Setser A, Basch E. Patient-reported outcomes and the evolution of adverse event reporting in oncology. Journal of Clinical Oncology. 2007;25(32):5121-7.
221. J M. Tracing books as patient-held files. Available from: http://web.uct.ac.za/depts/cgc/Jonathan/TracingBookPoster.htm. [Accessed 12th December 2012].
222. Allen EC, C.I.R. Mandimika, N. Barnes, K. Eliciting adverse effects data from participants in clinical trials (Protocol). The Cochrane Library: John Wiley & Sons; 2013.
223. Karita E, Ketter N, Price MA, Kayitenkore K, Kaleebu P, Nanvubya A, et al. CLSI-derived hematology and biochemistry reference intervals for healthy adults in eastern and southern Africa. PloS One. 2009;4(2):e4401.
224. Aronson JK, Ferner RE. Clarification of terminology in drug safety. Drug Safety. 2005;28(10):851-70.
225. Vandenbroucke JP. Benefits and harms of drug treatments. BMJ (Clinical research ed). 2004;329(7456):2-3.
226. Chou R, Fu R, Carson S, Saha S, Helfand M. Methodological shortcomings predicted lower harm estimates in one of two sets of studies of clinical interventions. Journal of Clinical Epidemiology. 2007;60(1):18-28.
227. Eziefula AC, Gosling R, Hwang J, Hsiang MS, Bousema T, von Seidlein L, et al. Rationale for short course primaquine in Africa to interrupt malaria transmission. Malaria Journal. 2012;11:360.
228. Fegan GW, Lang TA. Could an open-source clinical trial data-management system be what we have all been looking for? PLoS Medicine. 2008;5(3):e6.
229. Loke YK PD, Herxheimer A. Adverse effects. In: Higgins JPT GS, editor. Cochrane Handbook for Systematic Reviews of Interventions The Cochrane Collaboration.
230. Johnson MO, Neilands TB. Neuroticism, side effects, and health perceptions among HIV-infected individuals on antiretroviral medications. Journal of Clinical Psychology in Medical Settings. 2007;14(1):69-77.
231. Nestoriuc Y, Orav EJ, Liang MH, Horne R, Barsky AJ. Prediction of nonspecific side effects in rheumatoid arthritis patients by beliefs about medicines. Arthritis Care & Research. 2010;62(6):791-9.
232. Brugge D, Edgar T, George K, Heung J, Laws MB. Beyond literacy and numeracy in patient provider communication: focus groups suggest roles for empowerment, provider attitude and language. BMC Public Health. 2009;9:354.

233. Sandhu H, Adams A, Singleton L, Clark-Carter D, Kidd J. The impact of gender dyads on doctor-patient communication: a systematic review. Patient Education and Counseling. 2009;76(3):348-55.
234. Ong LM, de Haes JC, Hoos AM, Lammes FB. Doctor-patient communication: a review of the literature. Social Science & Medicine (1982). 1995;40(7):903-18.
235. M S. Multinational clinical trials: breaking language and cultural barriers. Applied Clinical Trials. 2003.
236. L C. The influence of the nocebo effect in clinical trials. Open Access Journal of Clinical Trials. 2012;4:61-8.
237. QSR International. Available from: http://www.qsrinternational.com/.
238. Auerbach CF SL. Qualitative data: an introduction to coding and analysis. New York: New York University Press; 2003. Accessed 15th August 2015.
239. Lunet N, Bastos J, Cumaio F, Silva P, Dias E, Barros H. Recall of drug utilization depends on subtle structural questionnaire characteristics. Pharmacy World & Science : PWS. 2008;30(2):175-81.
240. Bowling A. Mode of questionnaire administration can have serious effects on data quality. Journal of Public Health. 2005;27(3):281-91.
241. Rose N NC. Biological citizenship. In: A O, editor. Global Assemblages: technology, politics and ethical anthropological problems Oxford, UK: Blackwell Publishing Ltd; 2005. p. 439–63.
242. Molyneux CS, Peshu N, Marsh K. Understanding of informed consent in a low-income setting: three case studies from the Kenyan coast. Social Science & Medicine (1982). 2004;59(12):2547-59. Epub 2004/10/12.
243. Chandler CI, Hall-Clifford R, Asaph T, Pascal M, Clarke S, Mbonye AK. Introducing malaria rapid diagnostic tests at registered drug shops in Uganda: limitations of diagnostic testing in the reality of diagnosis. Social Science & Medicine (1982). 2011;72(6):937-44.
244. Charles C, Whelan T, Gafni A. What do we mean by partnership in making decisions about treatment? BMJ (Clinical Research ed). 1999;319(7212):780-2.
245. Achmat Z, Simcock J. Combining prevention, treatment and care: lessons from South Africa. AIDS (London, England). 2007;21 Suppl 4:S11-20.
246. Barlow F, Scott C, Coghlan B, Lee P, White P, Lewith GT, et al. How the psychosocial context of clinical trials differs from usual care: a qualitative study of acupuncture patients. BMC Medical Research Methodology. 2011;11:79.
247. Helman C. Culture, health and illness. 5 ed. London: Hodder Arnold; 2007.
248. Goodyear-Smith F, Buetow S. Power issues in the doctor-patient relationship. Health care analysis. 2001;9(4):449-62.
249. Hesse-Biber S. Weaving a multimethodology and mixed methods praxis into randomized control trials to enhance credibility. Qualitative Inquiry. 2012;18(10):876-89.
250. Talisuna AO, Staedke SG, D'Alessandro U. Pharmacovigilance of antimalarial treatment in Africa: is it possible? Malaria Journal. 2006;5:50.
251. National policy on traditional medicine and regulation of herbal medicines: report of a WHO global survey. Geneva: World Health Organization, 2005.
252. Conover EA. Over-the-counter products: nonprescription medications, nutraceuticals, and herbal agents. Clinical Obstetrics and Gynecology. 2002;45(1):89-98.
253. Abasiubong F, Bassey EA, Udobang JA, Akinbami OS, Udoh SB, Idung AU. Self-Medication: potential risks and hazards among pregnant women in Uyo, Nigeria. The Pan African Medical Journal. 2012;13:15.
254. Christianson A HC, Modell B. March of Dimes Global Report on Birth Defects. White Plains, New York: March of Dimes Birth Defects Foundation, 2006.

255. Dellicour S, Brasseur P, Thorn P, Gaye O, Olliaro P, Badiane M, et al. Probabilistic record linkage for monitoring the safety of artemisinin-based combination therapy in the first trimester of pregnancy in Senegal. Drug Safety. 2013;36(7):505-13.
256. Peltzer K, Phaswana-Mafuya N, Treger L. Use of traditional and complementary health practices in prenatal, delivery and postnatal care in the context of HIV transmission from mother to child (PMTCT) in the Eastern Cape, South Africa. African journal of traditional, complementary, and alternative medicines. 2009;6(2):155-62.
257. Global Health Observatory Data Repository. World Health Organisation; Available from: http://apps.who.int/gho/data/node.country.regionAFR. [Accessed 15th August 2015].
258. Dawson S ML, Tallo VL. A manual for the use of focus groups. 1993.
259. Finlayson K, Downe S. Why do women not use antenatal services in low- and middle-income countries? A meta-synthesis of qualitative studies. PLoS Medicine. 2013;10(1):e1001373.
260. Pell C, Menaca A, Were F, Afrah NA, Chatio S, Manda-Taylor L, et al. Factors affecting antenatal care attendance: results from qualitative studies in Ghana, Kenya and Malawi. PloS One. 2013;8(1):e53747.
261. Simkhada B, Teijlingen ER, Porter M, Simkhada P. Factors affecting the utilization of antenatal care in developing countries: systematic review of the literature. Journal of Advanced Nursing. 2008;61(3):244-60.
262. Mbonye AK, Neema S, Magnussen P. Treatment-seeking practices for malaria in pregnancy among rural women in Mukono district, Uganda. Journal of Biosocial Science. 2006;38(2):221-37.
263. Atuyambe L, Mirembe F, Tumwesigye NM, Annika J, Kirumira EK, Faxelid E. Adolescent and adult first time mothers' health seeking practices during pregnancy and early motherhood in Wakiso district, central Uganda. Reproductive Health. 2008;5:13.
264. Respectful maternity care toolkit The Knowledge for Health (K4Health) Project, Johns Hopkins Bloomberg School of Public Health's Center for Communication Programs.
265. Broussard CS, Louik C, Honein MA, Mitchell AA, National Birth Defects Prevention S. Herbal use before and during pregnancy. American Journal of Obstetrics and Gynecology. 2010;202(5):443 e1-6.
266. Adusi-Poku Y, Edusei AK, Bonney AA, Tagbor H, Nakua E, Otupiri E. Pregnant women and alcohol use in the Bosomtwe district of the Ashanti region-Ghana. African Journal of Reproductive Health. 2012;16(1):55-60.
267. Namagembe I, Jackson LW, Zullo MD, Frank SH, Byamugisha JK, Sethi AK. Consumption of alcoholic beverages among pregnant urban Ugandan women. Maternal and Child Health Journal. 2010;14(4):492-500.
268. Petersen Z, Nilsson M, Everett K, Emmelin M. Possibilities for transparency and trust in the communication between midwives and pregnant women: the case of smoking. Midwifery. 2009;25(4):382-91.
269. Grimaldi-Bensouda L, Rossignol M, Aubrun E, Benichou J, Abenhaim L, the PSG. Agreement between patients' self-report and physicians' prescriptions on nonsteroidal anti-inflammatory drugs and other drugs used in musculoskeletal disorders: the international Pharmacoepidemiologic General Research eXtension database. Pharmacoepidemiology and Drug Safety. 2012;21(7):753-9.
270. Gallo M, Einarson A, Koren G. Herbal medicine use in pregnancy: a new frontier in clinical teratology. Birth defects research Part B, Developmental and Reproductive Toxicology. 2003;68(6):499-500.
271. McIlleron H, Abdel-Rahman S, Dave JA, Blockman M, Owen A. Special populations and pharmacogenetic issues in tuberculosis drug development and clinical research. The Journal of Infectious Diseases. 2015;211 Suppl 3:S115-25.

272. McCambridge J, Kypri K, Elbourne D. Research participation effects: a skeleton in the methodological cupboard. Journal of Clinical Epidemiology. 2014;67(8):845-9.
273. Mothupi MC. Use of herbal medicine during pregnancy among women with access to public healthcare in Nairobi, Kenya: a cross-sectional survey. BMC Complementary and Alternative Medicine. 2014;14:432.
274. Kaingu CK, Oduma JA, Kanui TI. Practices of traditional birth attendants in Machakos District, Kenya. Journal of Ethnopharmacology. 2011;137(1):495-502.
275. Hill J, D'Mello-Guyett L, Hoyt J, van Eijk AM, ter Kuile FO, Webster J. Women's access and provider practices for the case management of malaria during pregnancy: a systematic review and meta-analysis. PLoS Medicine. 2014;11(8):e1001688.
276. Geissler PW. 'TRANSPORT TO WHERE?'. Journal of Cultural Economy. 2011;4(1):45-64.
277. Declaration of Helsinki - Ethical Principles for Medical Research Involving Human Subjects. World Medical Association, 64th WMA General Assembly, Fortaleza, Brazil, 2013.
278. Wiseman V, Conteh L, Matovu F. Using diaries to collect data in resource-poor settings: questions on design and implementation. Health Policy and Planning. 2005;20(6):394-404.
279. Chandler CIR, Reynolds, J., Palmer, J. J., & Hutchinson, E. ACT Consortium guidance: Qualitative Methods for International Health Intervention Research. 2013. Available from: Available at www.actconsortium.org/qualitativemethodsguidance. [Accessed 15th August 2015].
280. Tates K, Elbers E, Meeuwesen L, Bensing J. Doctor-parent-child relationships: a 'pas de trois'. Patient Education and Counseling. 2002;48(1):5-14.
281. Brasseur P, Vaillant MT, Olliaro PL. Anti-malarial drug safety information obtained through routine monitoring in a rural district of South-Western Senegal. Malaria Journal. 2012;11:402.
282. Lassere MN, Johnson KR, Woodworth TG, Furst DE, Fries JF, Kirwan JR, et al. Challenges and progress in adverse event ascertainment and reporting in clinical trials. The Journal of Rheumatology. 2005;32(10):2030-2.
283. Gargon E, Williamson PR, Altman DG, Blazeby JM, Clarke M. The COMET Initiative database: progress and activities from 2011 to 2013. Trials. 2014;15:279.
284. Kabudula CW, Clark BD, Gomez-Olive FX, Tollman S, Menken J, Reniers G. The promise of record linkage for assessing the uptake of health services in resource constrained settings: a pilot study from South Africa. BMC Medical Research Methodology. 2014;14:71. Epub 2014/06/03.
285. Lasch KE, Marquis P, Vigneux M, Abetz L, Arnould B, Bayliss M, et al. PRO development: rigorous qualitative research as the crucial foundation. Quality of Life Research: an international journal of quality of life aspects of treatment, care and rehabilitation. 2010;19(8):1087-96.
286. Cox K. Assessing the quality of life of patients in phase I and II anti-cancer drug trials: interviews versus questionnaires. Social Science & Medicine (1982). 2003;56(5):921-34
287. Baum F. Researching public health: behind the qualitative-quantitative methodological debate. Social Science & Medicine (1982). 1995;40(4):459-68.
288. Botelho EM, Elstad EA, Taubenberger SP, Tennstedt SL. Moderating perceptions of bother reports by individuals experiencing lower urinary tract symptoms. Qualitative Health Research. 2011;21(9):1229-38.
289. Justice AC, Chang CH, Rabeneck L, Zackin R. Clinical importance of provider-reported HIV symptoms compared with patient-report. Medical Care. 2001;39(4):397-408.

290. Dickert N, Sugarman J. Ethical goals of community consultation in research. American Journal of Public Health. 2005;95(7):1123-7.
291. Gikonyo C, Bejon P, Marsh V, Molyneux S. Taking social relationships seriously: lessons learned from the informed consent practices of a vaccine trial on the Kenyan Coast. Social Science & Medicine (1982). 2008;67(5):708-20.
292. Geissler PW, Kelly A, Imoukhuede B, Pool R. 'He is now like a brother, I can even give him some blood'--relational ethics and material exchanges in a malaria vaccine 'trial community' in The Gambia. Social Science & Medicine (1982). 2008;67(5):696-707.
293. Marsh VM, Kamuya DK, Parker MJ, Molyneux CS. Working with concepts: the role of community in international collaborative biomedical research. Public Health Ethics. 2011;4(1):26-39.
294. Forsythe LP, Ellis LE, Edmundson L, Sabharwal R, Rein A, Konopka K, et al. Patient and stakeholder engagement in the PCORI pilot projects: description and lessons learned. Journal of General Internal Medicine. 2015.
295. Svensson S, Kjellgren KI. Adverse events and patients' perceptions of antihypertensive drug effectiveness. J Hum Hypertens. 2003;17(10):671-5.
296. Bukirwa H, Nayiga S, Lubanga R, Mwebaza N, Chandler C, Hopkins H, et al. Pharmacovigilance of antimalarial treatment in Uganda: community perceptions and suggestions for reporting adverse events. Tropical Medicine & International Health 2008;13(9):1143-52.
297. Kamat VR, Nyato DJ. Community response to artemisinin-based combination therapy for childhood malaria: a case study from Dar es Salaam, Tanzania. Malaria Journal. 2010;9:61.
298. Mangesho PE, Reynolds J, Lemnge M, Vestergaard LS, Chandler CI. "Every drug goes to treat its own disease..." - a qualitative study of perceptions and experiences of taking anti-retrovirals concomitantly with anti-malarials among those affected by HIV and malaria in Tanzania. Malaria Journal. 2014;13:491.
299. Kabanywanyi AM, Mulure N, Migoha C, Malila A, Lengeler C, Schlienger R, et al. Experience of safety monitoring in the context of a prospective observational study of artemether-lumefantrine in rural Tanzania: lessons learned for pharmacovigilance reporting. Malaria Journal. 2010;9:205.
300. Adedeji AA, Sanusi B, Tella A, Akinsanya M, Ojo O, Akinwunmi MO, et al. Exposure to anti-malarial drugs and monitoring of adverse drug reactions using toll-free mobile phone calls in private retail sector in Sagamu, Nigeria: implications for pharmacovigilance. Malaria Journal. 2011;10:230.

Supplementary Table S3.1: Characteristics of included studies (Chapter 3)

Allen 2013

Methods	Methods study nested in trial. Subset of adults with or without malaria enrolled in 2 open-label artemether-lumefantrine/antiretroviral interaction trials (South Africa, N=16; Tanzania, N=76). Within-participant comparison of cumulative data elicited by 2 consecutive methods prior to treatment and after 3-7 days. Third method then applied with participants who reported differently between first 2 methods (South Africa, N=11; Tanzania, N=16). Participants' experiences of illness and treatment and reporting behaviour also explored qualitatively.
Data	Treatment-emergent AEs, severity assessed by investigator. Also, previous and concomitant medications, medical histories.
Comparisons	ELICITATION METHOD 1 General open-ended verbal enquiry about health and medicine use without reference to particular conditions, body systems or treatments. ELICITATION METHOD 2 Immediately after Method 1: verbal enquiry with reference to checklists of potential health issues and medicines. ELICITATION METHOD 3 Within 7 days of Methods 1 and 2: in-depth interview: prompted narrative of the participant's trial experience, reflection on previous ill health and medicines used and photographs of typical over-the-counter and traditional medicines available to the study populations.
Outcomes	Number (%) additional AEs, medications and medical histories by previous method. AE severity description. Themes, theoretical interpretation of participants' experiences related to differential reporting between methods. Cannot distinguish between treatments and not informative to make a direct comparison between sites due to differences in the participant populations and trial designs.
Notes	A majority of fields in the checklists were common to both trials although they could not be harmonised fully. Answers probed according to common clinical practice in eliciting a medical and treatment history.

Avery 1967

Methods	Methods study nested in trial. Subset of depressed inpatients enrolled in pilot study of chlorpromazine with or without procyclidine versus placebo (US, N=23). Between-participant comparison of data elicited by 2 randomly

	allocated methods prior to treatment and at weekly intervals for 5 weeks.
Data	Symptoms (unclear if treatment-emergent). Severity estimated by weighting the gross symptoms by a factor derived from degree of subjective discomfort (0 to 3).
Comparisons	ELICITATION METHOD 1 Verbal enquiry "Have you noticed any change in bodily function or had any physical complaints in the past week?". ELICITATION METHOD 2 Method 1 plus specific questions from a study checklist of possible drug side effects. The question in elicitation method 1 plus specific questions from a checklist of possible drug side effects.
Outcomes	Means and ranges of number and severity of symptoms by method and treatment (p=0.05, one-tailed test of significance using Mann-Whitney U test).
Notes	Method 2: National Institute of Mental Health study checklist of possible drug side effects (NB assumed not all were asked).

Barber 1995

Methods	Methods study nested in trial. Subset of adults with ocular hypertension or open-angle glaucoma randomised to 2% dorzolamide or 2% pilocarpine plus 0.5% timolol in a cross-over trial (US, N=47 [pilocarpine phase only reported due to lack of AEs with dorzolamide]). Within-participant comparison of data elicited by 2 concurrent methods 1) prior to treatment, days 14, 30 and 2) throughout trial.
Data	Adverse experiences (unclear if treatment-emergent), participant assessment of bothersomeness (6-point scale), quality of life.
Comparisons	ELICITATION METHOD 1 Interviewer-administered questionnaire (COMTOL). NB domain scores calculated from average of symptoms. ELICITATION METHOD 2 Participants instructed to call investigator if they experienced an AE and investigator asked participants at each visit if they had experienced any AE since their last visit (not clear which time point in relation to Method 1).
Outcomes	Number (%), mean scores (SD) AE frequency and bother domain scores for those who did not report any spontaneous AEs but indicated on

	questionnaire they had the AE. Relationship between AEs and quality of life.
Notes	Method 1: validated questionnaire that captures the frequency and bother of common side effects (i.e. ocular and other local effects, and effects on visual function) of topical therapy for lowering intraocular pressure. In addition, the questionnaire measures the extent to which these side effects and any associated limitations in routine living activities interfere with health-related quality of life, medication compliance, and patient satisfaction with the medication.

Barrowman 1970

Methods	Methods study nested in class experiment. Healthy medical students administered pentagastrin 6 µg/kg or 0.9% sodium chloride (UK, N=24). Within-participant comparison of data elicited by 2 consecutive methods over 10 minutes post treatment.
Data	Unwanted subjective effects/symptoms (assumed treatment-emergent).
Comparisons	ELICITATION METHOD 1 During first 9 minutes post-dose observer, on 3 occasions, asked subject to describe any unusual sensation. ELICITATION METHOD 2 10 minutes post-dose subject asked directly about certain symptoms.
Outcomes	Number of symptoms by method and treatment.
Notes	Method 2: Symptoms known or suspected to occur after pentagastrin, and some control items (headache, dryness of mouth, increased salivation). Instructions given to the observer about the questioning methods, including timing and how to complete the data elicited according to each method.

Bent 2006

Methods	Methods study nested in trial. Subset of healthy men with benign prostatic hyperplasia enrolled in a trial of 'saw palmetto' (US, N=214) . Between-participant comparison of data elicited by 1 of 3 randomly allocated methods at the end of a 1-month, single-blind, placebo run-in period.
Data	AEs (unclear if treatment-emergent), seriousness assessed by investigator.

Comparisons	ELICITATION METHOD 1 Self-administered open-ended question ("Did you have any significant medical problem since the last study visit?"). If "yes", participants asked to identify medical problem (recorded by study assistant on same checklist as Method 3 group). ELICITATION METHOD 2 Self-administered open-ended explicit question ("Since the last study visit, have you limited your usual daily activities for more than 1 day because of a medical problem?"). If "yes", participants asked to identify medical problem (recorded by study assistant on same checklist as Method 3 group). ELICITATION METHOD 3 Self-administered checklist ("Since the last visit, have you experienced any of the following?": 53 symptoms, grouped by anatomical region).
Outcomes	Number/type of AEs by method. Difference in proportion of participants reporting ≥ 1 AE by method ($\chi 2$). SAE description.
Notes	Method 3: checklist developed after a unpublished review of checklists used in earlier clinical trials at the same institution. Self-completed, although a study assistant recorded medical problems on the checklist.

Borghi 1984

Methods	Methods study nested in trial. Adult antihypertensive outpatients enrolled in a multi-centre, double-blind, randomised cross-over trial of oxprenolol versus chlorthiazide with single-blind placebo wash-out periods (Italy, N=223/227). Between-participant comparison of data elicited by a conventional approach followed by 1 of 2 randomly assigned methods throughout the trial.
Data	Unwanted effects (unclear if treatment-emergent).
Comparisons	ELICITATION METHOD 1 Reported signs and symptoms evaluated by physician (suggests filtering of reports depending on judgement about causality so not data included in review). ELICITATION METHOD 2 Method 1 plus self-completed checklist of 49 items requiring yes/no (sequence changed each visit).

	ELICITATION METHOD 3 Method 1 plus self-completed blank card, same format as Method 2, for participant to report signs and symptoms experienced.
Outcomes	Number of participants (%) with ≥ 1 unwanted effect by treatment.
Notes	Method 2: 49 items consisting of pharmacological unwanted effects linked to the most common antihypertensive drugs, mainly ß-blockers and diuretics. The investigator was neither informed of the results of the questionnaires, nor did they help the participant to fill them in, so as not to influence the data collection.

Brent 2009

Methods	Methods integral to trial. Adolescent outpatients with moderate to severe depressive disorder and taking an SSRI randomised to another SSRI or venlafaxine with or without cognitive behaviour therapy (US, N=334). Between-participant comparison of data elicited by 2 non-randomly allocated methods at each visit over 12 weeks.
Data	Self-harm AEs (suicidal and non-suicidal self-injury), assumed treatment-emergent, seriousness assessed by investigator.
Comparisons	ELICITATION METHOD 1 Spontaneous report of self-harm (no details). ELICITATION METHOD 2 Weekly monitoring using Brief Suicide Severity Rating Scale: 1) rating of suicidal ideation 0 to 5 and 2) rating of suicidal behaviour 0 to 5 using Columbia Classification Algorithm of Suicide Assessment; two-point change on either scale determined if a suicidal AE occurred.
Outcomes	Proportion of AEs by method (standard univariate statistics). Times to event per method (Kaplan-Meier). AE vs SAE description.
Notes	Method 2 involved standard validated instruments: Brief Suicide Severity Rating Scale and Columbia Classification Algorithm of Suicide Assessment are published tools

Ciccolunghi 1975

Methods	Methods study outside of trial. Adult employees of research company (clinical research and production departments). Between-participant

	comparison of data elicited by 1 of 3 randomly allocated methods on one occasion by group, healthy versus those taking medication (Switzerland, N=416).
Data	Symptoms (not necessarily treatment-emergent), participant assessment of severity (2 point scale)
Comparisons	ELICITATION METHOD 1 Open-ended questionnaire with 3 entry lines. ELICITATION METHOD 2 Open-ended questionnaire with 10 entry lines. ELICITATION METHOD 3 Checklist of 38 items.
Outcomes	Number of participants with ≥1 symptom (%), range, median by treatment (healthy versus medication). Severity description.
Notes	Methods distributed by internal mail with addressed envelope for return. Anonymity assured.

de Vries 2013

Methods	Tool validation outside of trial. Subset of adult outpatients dispensed an oral glucose lowering drug (Netherlands, N=90). Between-participant comparison of data elicited by 2 randomly allocated methods. Description of feasibility including self-reported time to completion and ease of use (5-point Likert).
Data	Adverse drug events (ADEs) (not necessarily treatment-emergent). Feasibility of completion.
Comparisons	ELICITATION METHOD 1 Email invite for internet-based self-administered questionnaire with ADEs categorized in 16 body categories (T1), repeat after 1 week with no body categories (T2). ELICITATION METHOD 2 Email invite for internet-based self-administered questionnaire with ADEs not categorized in 16 body categories (T1), repeat after 1 week with body categories (T2).

Outcomes	Number of ADEs by method (χ2, Mann-Whitney U tests). Agreement of methods (Z value). Description of feasibility outcomes.
Notes	Content validation of common ADEs drafted in layman terms with reference to CTCAE v4.0, existing symptom and ADE checklists, then coded, categorised into body categories using MedDRA® classifications. Open-ended option for "other", questions relating to duration, frequency, seriousness and causality based on literature.

de Vries 2014

Methods	Tool validation outside of trial. Subset of adult outpatients dispensed an oral glucose lowering drug (Netherlands, N=78). Between-participant comparison of data elicited by 2 consecutive methods (random allocation of second method using a 4-week or 3-month recall period).
Data	Adverse drug events (ADEs) (not necessarily treatment-emergent), participants' assessment of nature and causality.
Comparisons	ELICITATION METHOD 1 Gold standard: paper-based daily diary completed for 3 months: an open-ended question asking for symptoms experienced and closed-ened question about attribution to any drug taken. ELICITATION METHOD 2 Email invite for internet-based self-administered questionnaire with ADEs (symptoms in lay terms) relating to past 4 weeks. ELICITATION METHOD 3 Email invite for internet-based self-administered questionnaire with ADEs (symptoms in lay terms) relating to past 3 months.
Outcomes	Sensitivities and positive predictor values (CI) at class and specific ADE levels.
Notes	Content validation of common ADEs drafted in layman terms with reference to CTCAE v4.0, existing symptom and ADE checklists, then coded, categorised into body categories using MedDRA® classifications. Open-ended option for "other", questions relating to duration, frequency, seriousness and causality based on literature.

Downing 1970

Methods	Methods study nested in trial. Adults with mild to moderate anxiety and/or depression receiving one of several antidepressants (amitriptyline, iprindole), tranquillizers (chlordiazepoxide, diazepam, fluphenazine) or placebo in double-blind trials (US, N=123) . Within-participant comparison of data elicited by 2 consecutive methods 4 weeks post-treatment.
Data	Side reactions/effects, participant assessment of intensity, discomfort and opinion on relationship to study drug. NB only those symptoms that the participant related to medication reported.
Comparisons	ELICITATION METHOD 1 Open-ended question (O) week 2 and 4: "How are you feeling?". If no reference to drug-related symptomology: "How else are you feeling?" then "How does the drug make you feel?". ELICITATION METHOD 2 Structured (S) question 28-item questionnaire as basis of structured interview at week 4 after Method 1. If AE report, patient asked to estimate intensity (3 point scale), discomfort (4 point scale) and whether symptom was felt due to study medicine.
Outcomes	Incidence of side reactions (0 or ≥ 1) per method. Comparative incidence between methods ($\chi 2$ using McNemar's formula, $p < 0.01$). Number of side effects per participant per treatment. Number of new side effects by method 2. Mean intensity and discomfort scores per participant (average all drug-related symptoms reported). Number (%) of events attributed or not to treatment.
Notes	Method 2: 23 common medication effects and 5 highly unlikely to be related. The latter were captured in a miscellaneous category on coding sheet. Methods applied by the treating physician (extensive training in interviewing and rating procedures). Reports entered onto data sheet with categories provided for medication-produced disturbances frequently associated with the medications. Symptoms unlikely to be associated with the medications recorded as miscellaneous.

Greenhill 2004

Methods	Methods study outside of trial. Children initiating treatment with 1 or more psychotropic medicines in the past 60 days and attending outpatient visits (US, N=59). Within-participant comparison of data elicited by 3 consecutive sections of an instrument delivered as a scripted interview at a routine follow-up visit.
Data	AEs (any unfavourable event that occurs during treatment or a clinical trial, regardless of cause), severity and clinical relevance assessed by

	investigator.
Comparisons	ELICITATION METHOD 1 3-question general inquiry (GI): "Has ____ had any physical or health problems since….? I'm talking about something that started to become a problem during this time or an old problem that got much worse.", "Have there been activities that ____ didn't do as often or that he/she didn't do at all because of not feeling well since _____?", Since ____, has ____ said that his/her body feels funny... or that he/she has any aches or pains... or that some part of him/her hurts or doesn't feel well?". ELICITATION METHOD 2 After Method 1: drug-specific inquiry (DSI) - 18 questions about clinically important AEs for various medicines. ELICITATION METHOD 3 After Method 1 and 2: body system review (BSR) - 24 questions.
Outcomes	Number (%) of AEs first elicited by method (by AE severity and clinical relevance). Time for administration by method.
Notes	Methods 1 and 2 (same instrument): semi-structured interview (SMURF) constructed from an existing instrument (SAFTEE) as a scripted interview, with instructions, demographic queries and a glossary of preferred AE terms. Only AEs where it would be malpractice not to assess were included. Experienced clinicians (95% psychiatrists and 5% nurses) trained to evaluate and treat children in child psychiatric settings conducted the interviews. They received 1 hour of telephone training in the administration of the SMURF. AEs elicited were captured on another form using SAFTEE preferred terms and related details.

Hermans 1994

Methods	Methods integral to trial. Adults with mild to moderate hypertension enrolled in a double-blind, randomised trial comparing isradipine and amlodipine (Belgium, N=205). Within-participant comparison of data elicited by 2 consecutive methods at baseline and after 6 weeks.
Data	AEs, order experienced.
Comparisons	ELICITATION METHOD 1 Verbal enquiry 'How have you felt since your last visit?'. ELICITATION METHOD 2

	Self-completed written questionnaire: swollen ankles, headache, flushing, palpitations, dizziness or nausea.
Outcomes	Number (%) of AEs by method. Incidence of participants with AE by method ($\chi 2$).
Notes	Method 2: anticipated side-effects of dihydropyridine calcium antagonists. Assumed cardiologists and nephrologists asking questions.

Huskisson 1974

Methods	Methods study nested in trial. Participants with rheumatoid arthritis enrolled in a RCT of aspirin versus fenoprofen (UK, N=60). Between-group non-random comparison of data elicited by 1 or 2 methods over 6 months (timing of assessments unknown).
Data	Side effects, severity (not clear who assessed severity).
Comparisons	ELICITATION METHOD 1 Verbal enquiry: "Have you noticed any new symptoms which might be related to the treatment?". ELICITATION METHOD 2 Method 1 plus check list of 21 possible side effects recorded as absent, slight, moderate, or severe (0, 1, 2, or 3); side effect score from sum of values.
Outcomes	% AE side effect scores by method and treatment (sum of severity 0 [absent], 1, 2, 3 x 100/number of participants). Cross-tabulation by method. NB those side effects that were not significantly different between treatments (by either method) were grouped in analysis as 'irrelevant'.
Notes	Method 2: tinnitus, deafness, gastrointestinal complaints and others with no obvious relevance and others with no obvious relevance.

Jacobson 1987

Methods	Tool validation within trial. Adults with schizophrenia, major depressive episode with psychotic features, moderate or greater anxiety, depression or insomnia enrolled in inpatient or outpatient trials investigating drug treatments (US, N=134). Within-participant comparison of data elicited by 2 consecutive methods within a structure verbal interview at weekly trial

	visits.
Data	AEs.
Comparisons	ELICITATION METHOD 1 General inquiry (GI): "Have you had any physical or health problems during the past week [or specified assessment interval]? Have you noticed any changes in your physical appearance during the past week [or specified assessment interval]? Have you cut down on the things you usually do because of not feeling well physically during the past week [or specified assessment interval]?". ELICITATION METHOD 2 After Method 1: systematic inquiry (SI) which is the GI plus a review of 23 body systems.
Outcomes	Number of AEs by method, mean number of AEs per assessment, mean severity and impairment.
Notes	SAFTEE developed by the National Institute of Mental Health. This study part of validation exercise. NB participants were randomly assigned to different staff applying each elicitation method.

Kruft 2007

Methods	Methods integral to trial. Retrospective meta-analysis of 4 double-masked, randomised, cross-over ophthalmic trials involving various drugs or placebo (multinational, N=223). Within-participant comparison of data elicited by 2 methods.
Data	Ocular AEs.
Comparisons	ELICITATION METHOD 1 General query: "How are you doing since your last visit?". ELICITATION METHOD 2 Solicited ophthalmic symptom query checklist, including visual analogue scales (VAS).
Outcomes	Number (%) ophthalmic symptoms by method.
Notes	Method 2: some VAS in trials were validated instruments, not clear if this one was.

Landén 2005

Methods	Methods study nested in trial. Adults with treatment-refractory depression enrolled in a placebo-controlled RCT of bupirone-augmentation of SSRI therapy (Sweden and Norway, N=119). Within-participant comparison of data elicited by 2 consecutive methods before and 4 weeks post-treatment.
Data	Sexual side-effects.
Comparisons	ELICITATION METHOD 1 Non-leading question such as: "Have you felt different in any way since you started the new treatment?". ELICITATION METHOD 2 After Method 1: direct questions from UKU side-effect rating scale (none, mild, moderate, severe) for 3 symptoms of sexual dysfunction.
Outcomes	Number sexual side effects by method, OR (Pearson $\chi 2$, Yates correction as appropriate).
Notes	Method 2: UKU is a validated instrument.

Lundberg 1980

Methods	Methods study nested in trial. Adult male healthy volunteers enrolled in a double-blind, placebo-controlled, 2-factor cross-over trial of diphenhydramine versus terfenadine (US, N=12). Within-participant comparison of data elicited by 2 consecutive methods on 2 occasions at 3 visits over 9 days.
Data	Side-effects.
Comparisons	ELICITATION METHOD 1 Somesthetic Inventory (self-completed): 54 body feelings, 9 fillers assessed through a visual analogue scale (VAS). ELICITATION METHOD 2 After Method 1: self-completed Side Effects Report of 24 terms assessed with VAS.
Outcomes	2-factor, repeated-measures analysis of variance for body feeling or side-effect reported by at least 50% of sample.

Notes	Method 1: compilation and organization of data on side-effects of antihistamines (18 terms also in Method 1). Method 2: arbitrarily selected terms. Method 1: participants asked to 'attune to inner stimuli', close eyes and determine how body felt. Method 2: definitions provided to participants.

Monteiro 1987

Methods	Methods study nested in trial. Adults with severe OCD and DSM-III enrolled in a double-blind RCT of clomipramine vs placebo (UK, N=33/46). Within-participant comparison of data elicited by 3 methods.
Data	Sexual side-effects.
Comparisons	ELICITATION METHOD 1 Spontaneous reports at any time and at week 8. ELICITATION METHOD 2 Self-rated physical symptom questionnaire (including sexual function items) at weeks 0 and 8. ELICITATION METHOD 3 Structured interview enquiry (not clear when conducted assume at end of study).
Outcomes	Number (%) of participants reporting any sexual dysfunction by method. Could not split by trial arm.
Notes	

Nicholls 1980

Methods	Methods integral to trial. Adults with mild to moderate essential hypertension enrolled in a double-blind, double-dummy cross-over RCT of labetalol vs propranolol (UK, N=24). Within-participant comparison of data elicited by 2 consecutive methods at each visit and end of 8-week treatment period.
Data	Side-effects.
Comparisons	ELICITATION METHOD 1 Spontaneous/direct reporting (no details).

	ELICITATION METHOD 2 Self-administered questionnaire after assessment and the end of each treatment period.
Outcomes	% participants reporting each symptom by method and treatment arm. Average number of symptoms per patient per group.
Notes	

O'Connell 2007

Methods	Methods study nested in trial. Adolescent girls with dysmenorrhoea enrolled in a placebo-controlled RCT of ethinyl estradiol/levonorgestrel (US, N=76). Within-participant comparison of data elicited by 2 consecutive methods after a 3 month treatment period.
Data	Side effects, including depression.
Comparisons	ELICITATION METHOD 1 Open-ended question at one 3 month visit - participants asked to list any side effects or changes they experienced during the study. ELICITATION METHOD 2 After Method 1: participants asked if they experienced any of 12 specific side effects.
Outcomes	Number of AEs, % participants reporting ≥ 1 AE, median number of AEs by method and treatment arm.
Notes	Method 2: AEs commonly attributed to OCs, including headache, nausea, acne, abdominal pain, back pain, vomiting, breast tenderness, breast enlargement, mood swings, weight gain, premenstrual syndrome and irregular bleeding.

Os 1994

Methods	Methods integral to trial. Adults with mild to moderate hypertension enrolled in a double-blind, double-dummy RCT of lisinopril versus nifedipine (Norway, N=828). Within-participant comparison of data elicited by 3 consecutive methods several times during the trial (unclear order).
Data	Side-effect (cough).

Comparisons	ELICITATION METHOD 1 Spontaneous reporting (no details). ELICITATION METHOD 2 Direct questioning to be answered 'yes' or 'no' (no details). ELICITATION METHOD 3 Questionnaires consisting of VAS completed by participant and spouse independently.
Outcomes	Method 1: frequency (%), methods 2 and 3: cumulative incidence after 2, 6 and 10 weeks. Within-treatment changes by means of McNemar, between-treatment difference using log linear model. VAS between-treatment groups using ANOVA on ranks of changes from baseline.
Notes	Method 2: part of the ASPECT Scale - a tool for evaluation of 34 commonly experienced symptomatic side-effects of cardiovascular drugs.

Perez-Lloret 2012

Methods	Methods study outside trial. Adults with Parkinson's Disease and post-stroke controls (France, N=255) receiving at least 1 drug. Within-participant comparison of data elicited by 2 consecutive methods on 1 occasion.
Data	AEs (any untoward medical occurrence in a patient who is under any pharmacological treatment; the AE does not necessarily have to have a causal relationship with this treatment). Causality algorithm, intensity evaluated subjectively by trial staff.
Comparisons	ELICITATION METHOD 1 Verbal open enquiry: "Have you noticed any unpleasant effects of your medications during the previous week?". ELICITATION METHOD 2 After Method 1: verbal structured enquiry about the previous week using a pre-defined list of AEs.
Outcomes	Number of participants reporting ≥ 1 AE. Total number of AEs by method and population using $\chi 2$. Rate (%) of underreporting, 95% CI, binomial test for differences of AEs affecting > 10% of patients. Unpaired t-test/$\chi 2$ for comparing numerical or categorical variables, forward regression to identify independent factors related to spontaneous reporting.

Notes	Method 2: pre-defined list of most common ADRs to various anti-PD drugs from a literature search critically reviewed by a group of PD and pharmacovigilance specialist for consensus: general, GI, urinary, neuropsychiatric, dermatologic.

Rabkin 1992

Methods	Method nested within trial. Adults with bulimia, panic disorder, major depression or dysthymia enrolled in inpatient or outpatient trials (US, N=180/226) investigating drug treatments or placebo. Within-participant comparison of data elicited by 2 consecutive methods within a structure verbal interview pre-treatment and after 4 weeks.
Data	AEs
Comparisons	ELICITATION METHOD 1 General inquiry (GI): "Have you had any physical or health problems during the past week [or specified assessment interval]? Have you noticed any changes in your physical appearance during the past week [or specified assessment interval]? Have you cut down on the things you usually do because of not feeling well physically during the past week [or specified assessment interval]?". ELICITATION METHOD 2 GI plus Systematic inquiry (SI) which is a review of 23 body systems plus additional 11 items to represent side effects of MAOIs.
Outcomes	Number of AEs, type of AE, mean severity (removing comparative data from baseline), functional impairment, clinical action taken by method. Paired t-tests, OR with 95% CI (2-tailed).
Notes	SAFTEE developed by the National Institute of Mental Health. Additional 11 items research team's own choice. Severity subjectively graded and action taken noted. SAFTEE applied by either study psychiatrist or research nurse who had attended training meetings and had had a minimum of 3 practice audio-taped interviews reviewed by first author to assure adequacy of administration and rating.

Reilly 1992

Methods	Method nested within trial. Adults with mild to moderate essential hypertension enrolled in a multi centre, randomised, double-blind, placebo-controlled trial of clentiazem (US, N=92). Within-participant comparison of data elicited by 2 consecutive methods pre-treatment and

	at the final visit.
Data	Symptoms (not treatment-emergent), severity (bother on a 5-point scale from which 2 severity scales were created: symptom severity index checklist and symptom severity index open list. Scores calculated by multiplying no. of days bothered by 'extent bothered' for each symptom and summing scores). Secondary - arbitrary 20% change in SSI Checklist score used to represent clinically meaningful change to test a QoL instrument's responsiveness.
Comparisons	ELICITATION METHOD 1 Open-question: participants asked if they had any health-related symptoms or problems during the past 7 days. ELICITATION METHOD 2 Same questions with reference to a checklist immediately after Method 1.
Outcomes	% participants reporting ≥ 1 symptom by method. Methods compared for severity using Kruksal-Wallis Exact Test and Pearson product moment correlations.
Notes	Method 2: 24 symptoms associated with hypertension and anti-hypertensive therapy (previously used). The methods were applied 48 hours prior to the trial medical evaluation visit by a trained, full-time telephone interviewer. Need for further training was achieved though completed questionnaires being reviewed daily by a supervisor.

Rosenthal 1996

Methods	Methods integral to trial. Adults with mild to moderate essential hypertension enrolled in a double-blind RCT of quinapril versus metoprolol (Germany, N=5559). Within-participant comparison of data elicited by 2 consecutive methods.
Data	Adverse effects (those spontaneously-elicited were only presented if they matched questions on the questionnaire).
Comparisons	ELICITATION METHOD 1 Physician interview concerning general status of symptoms, development of signs (e.g. rash, swelling, bruises). ELICITATION METHOD 2 After Method 1: self-administered written questionnaire about signs and symptoms.

Outcomes	Number of AEs by method and treatment arm. Number (%) participants reporting AEs by method and treatment arm.
Notes	Method 2: formulated to elicit information concerning the appearance of signs and symptoms that could be related to ACE inhibitors or a beta-blocker, and any other symptoms that may reflect the patient's wellbeing. All spontaneously reported AEs were assigned to body systems categories according to COSTART criteria.

Sheftell 2004

Methods	Methods study outside trial. Adults with migraine aged ≤ 19 years taking triptans at 2 sites (US and Italy, N=415). Within-participant comparison of data elicited by 2 consecutive methods on one occasion.
Data	Adverse effects (those spontaneously-elicited were only presented if they matched questions on the questionnaire).
Comparisons	ELICITATION METHOD 1 Physician interview concerning general status of symptoms, development of signs (e.g. rash, swelling, bruises). ELICITATION METHOD 2 After Method 1: self-administered written questionnaire about signs and symptoms.
Outcomes	Number of AEs by method and treatment arm. Number (%) participants reporting AEs by method and treatment arm.
Notes	Method 2: formulated to elicit information concerning the appearance of signs and symptoms that could be related to ACE inhibitors or a beta-blocker, and any other symptoms that may reflect the patient's wellbeing. All spontaneously reported AEs were assigned to body systems categories according to Costart criteria.

Spilker 1987

Methods	Methods study outside of trial. Pharmacy staff, students and faculty. Between-participant comparison of data elicited by 1 of 2 randomly allocated methods on one occasion by group, healthy versus those taking medication (US, N=298).
Data	Symptoms (not treatment-emergent), participant assessment of severity.

Comparisons	ELICITATION METHOD 1 Self-completed questionnaires with 15 blank spaces to complete about demographics, tobacco and alcohol use, symptoms experienced in the past 72 hours, treatments used. ELICITATION METHOD 2 Self-completed questionnaire with a checklist of 25 symptoms about demographics, tobacco and alcohol use, symptoms experienced in the past 72 hours, treatments used.
Outcomes	Number symptoms/average no. of symptoms per person by method. Compared using T-tests.
Notes	Method 2: checklist used in a previous study. Handed out on a Thursday (middle of the week) to be completed and handed back in the same occasion.

Török 1984

Methods	Methods study nested in trial. Adults from 46 sites with hypertension, angina or arrhythmias enrolled in 3 trials (single arm or placebo controlled RCT) of chloranolol (Hungary, N=2066). Between-participant comparison of data elicited by 2 non-randomly allocated methods. Data only presented for participants taking chloranonol and for subjective GI related symptoms.
Data	AEs.
Comparisons	ELICITATION METHOD 1 Complaints reported spontaneously by the patients (in placebo-controlled trial only those symptoms during active drug phase recorded), or signs/symptoms observed by physician without using a list (objectively determined AEs not included in review). ELICITATION METHOD 2 As well as method 1, questions about side-effects listed in a questionnaire.
Outcomes	As objective signs also included, data here are only for GI symptoms, which are likely to be subjective - side-effects per 100 participants by method.
Notes	

Wallander 1991

Methods	Methods study nested in trial. Adults from 23 primary care sites with hypertension, angina or arrhythmias enrolled in a double-blind RCT of felodipine versus placebo added to metoprolol (Sweden, N=191/251). Within-participant comparison of data elicited by 3 consecutive methods at various visits up to 8 weeks.
Data	AEs.
Comparisons	ELICITATION METHOD 1 Complaint score (CS): "Have you had any of the following symptoms in the past month?" completed the day before baseline and final visit. ELICITATION METHOD 2 Subjective symptom assessment profile (SSAP): 41 item VAS completed day before the baseline and final visit.
Outcomes	Number of AEs by method. Bivariate relationships using Pitman's non-parametric permutation test, tests of paired data (before/after randomisation) using linear, non-parametric permutation, multivariate tests with Pitman's in multivariate form - reporting before randomisation treated as confounding variable.
Notes	Method 1: used in previous population studies, includes depression, tension, head, heart, lung, metabolism, musculoskeletal system, GI and urinary tracts. Patients place in an envelope and advised the physician would not have access. Method 2: validated instrument, highly correlated items in 6 domains, rest single items. After completion of Methods 1 and 2, they were put in an envelope and participants advised that the physician would not have access to the information. NB A previous method which involved question posed by a physician and then evaluated for association with the trial drugs was not included in the review comparison.

Wallin 1981

Methods	Methods study nested in trial. Patients with gonorrhoea enrolled in a study of bacampicillin (Sweden, N=515). Within-participant comparison of data elicited by 2 consecutive methods.
Data	Adverse reactions
Comparisons	ELICITATION METHOD 1 "Have you had any troubles from the drug?". ELICITATION METHOD 2 Checklist immediately after Method 1: "Have you noticed any of the

	following reactions: diarrhoea, nausea, vomiting, other gastrointestinal disturbances, skin eruptions, or other troubles?".
Outcomes	Number of additional AEs from previous method by method.
Notes	

Wernicke 2005

Methods	Methods integral to trials. Patients with various conditions enrolled in one of 3 double-blind, placebo-controlled RCTs of anonymous drugs (N=653: 219, 167, 267). Within-participant comparison of data elicited by 2 consecutive methods in participants who attended a visit where both methods were used.
Data	Treatment-emergent AEs.
Comparisons	ELICITATION METHOD 1 Unsolicited AEs by open-ended questioning; participants asked to report experiences since the last visit in their own words. ELICITATION METHOD 2 elicited AEs by standard questionnaires (Side Effects Checklist (child trial) BBAEQ-M (child/adolescent trial AMDP-5 (adult trial).
Outcomes	AEs reflecting same symptom in spontaneous and solicited methods selected. Ratio between rate reported by drug versus placebo (D/P) plotted for solicited on x-axis against spontaneous on y-axis. Also ratio of D/P ratios (Sp-So index): spontaneous D/P ratio divided by solicited D/P ratio (95% CI). Treatment differences compared using Fisher's exact test.
Notes	Method 2: Side Effects Checklist is based on the Subjective Treatment Emergent Symtpoms Scale (US National Institute of Mental health) - 30 items including general symptoms such as trouble sleeping, diarrhoea, headache, trouble eating. BBAEQ-M: 24 items rated 0-9, AMDP-5: 47 items rated 0-3. COSTART III was used to map actual terms to standard terms.

Yeo 1991

Methods	Methods integral to trial. Adults with hypertension enrolled in a placebo-controlled double-blind RCT comparing enalapril with nifedipine (ACE inhibitors) (UK, N=128). Within-participant comparison of data elicited by 2 methods (method 1 at each visit, method 2 pre-treatment, 8 and 24

	weeks/withdrawal).
Data	Side-effect (cough).
Comparisons	ELICITATION METHOD 1 "Have the tablets upset you in any way?" at each visit. ELICITATION METHOD 2 VAS at baseline, 8 weeks, 24 weeks/withdrawal (from "I never cough" to "I am always coughing").
Outcomes	Method 1: frequency of cough. Method 2: changes in mean scores and frequency of cough defined by an increase in VAS of ≥8mm. Between-treatment differences using Kolmogorov-Smirnov 2-sample test, $\chi 2$ with Yates correction, Fisher's exact, Wilcoxen rand sum tests.
Notes	

Supplementary Table S3.2: Characteristics of excluded studies (Chapter 3)

Author	Reasons for exclusion
Anderson RB, Testa MA. Symptom Distress Checklists as a Component of Quality of Life Measurement: Comparing Prompted Reports by Patient and Physician with Concurrent Adverse Event Reports via the Physician. Drug Information Journal 1994;28:89-114	No comparison of elicited data
Aspinall MB, Whittle J, Aspinall SL, Maher RL, Jr.Good CG. Improving adverse-drug-reaction reporting in ambulatory care clinics at a Veterans Affairs hospital. Am J Health Syst Pharm 2002;59(9):841-5	Lack of clinical trial focus
Atherton PJ, Burger KN, Loprinzi CL, Neben Wittich MA, Miller RC, Jatoi A, Sloan JA. Using the Skindex-16 and Common Terminology Criteria for Adverse Events to assess rash symptoms: results of a pooled-analysis (N0993). Support Care Cancer. 2012;20(8):1729-35	Included objective measure(s)
Basch E, Reeve BB, Mitchell SA, Clauser SB, Minasian LM, Dueck AC, Mendoza TR, Hay J, Atkinson TM, Abernethy AP, Bruner DW, Cleeland CS, Sloan JA, Chilukuri R, Baumgartner P, Denicoff A, St Germain D, O'Mara AM, Chen A, Kelaghan J, Bennett AV, Sit L, Rogak L, Barz A, Paul DB, Schrag D. Development of the National Cancer Institute's patient-reported outcomes version of the common terminology criteria for adverse events (PRO-CTCAE). J Natl Cancer Inst 2014;106(9):244	No comparison of elicited data
Bennett BK, Park SB, Lin CSY, Friedlander ML, Kiernan MC, Goldstein D. Impact of oxaliplatin-induced neuropathy: A patient perspective. Support Care Cancer 2012;20(11):2959-67	Assessed severity, experience of AEs already reported
van den Bergh JP, Bouts ME, van der Veer E, van der Velde RY, Janssen MJ, Geusens PP, Winkens B,Oldenhof NJ, van Geel TA. Comparing tolerability and efficacy of generic versus brand alendronate: a randomized clinical study in postmenopausal women with a recent fracture. PLoS One 2013;8(10):e78153	Not possible to compare data between methods
Bonierbale M, Lançon C, Tignol J. The ELIXIR study: evaluation of sexual dysfunction in 4557 depressed patients in France. Curr Med Res Opin 2003;19(2):114-24	Lack of clinical trial focus
Brown ES1, Jeffress J, Liggin JD, Garza M, Beard L. Switching outpatients with bipolar or schizoaffective disorders and substance abuse from their current antipsychotic to aripiprazole. J Clin Psychiatry 2005;66(6):756-60	Included objective measure(s)
Byerly MJ, Nakonezny PA, Fisher R, Magouirk B, Rush AJ. An empirical evaluation of the Arizona sexual experience scale and a simple one-item screening test for assessing antipsychotic-related sexual dysfunction in outpatients with schizophrenia and schizoaffective disorder. Schizophr Res 2006;81(2-3):311-6	Lack of clinical trial focus
Carreño M, Gil-Nagel A, Sánchez JC, Elices E, Serratosa JM, Salas-Puig J, Villanueva V, Porcel J. Strategies to detect adverse effects of antiepileptic drugs in clinical practice. Epilepsy Behav 2008;13(1):178-83	Lack of clinical trial focus
Coolbrandt, Van den Heede K, Vanhove E, De Bom A, Milisen K, Wildiers H. Immediate versus delayed self-reporting of symptoms and side effects during chemotherapy: does timing matter? Eur J Oncol Nurs 2011;15(2):130-6	Lack of clinical trial focus
Costa FV, Ambrosioni E, Magnani B. Side effects of antihypertensive drugs. Incidence and methods of collection. Int J Clin Pharmacol Biopharm 1979;17(9):405-9	Lack of clinical trial focus
De Smedt RH, Denig P, van der Meer K, Haaijer-Ruskamp FM, Jaarsma T. Self-reported adverse drug events and the role of illness perception and medication beliefs in ambulatory heart failure patients: A cross-	Lack of clinical trial focus

sectional survey. Int J Nurs Stud 2011;48(12):1540-50	
Downie FP, Mar Fan HG, Houédé-Tchen N, Yi Q, Tannock IF. Cognitive function, fatigue, and menopausal symptoms in breast cancer patients receiving adjuvant chemotherapy: evaluation with patient interview after formal assessment. Psychooncology 2006;15(10):921-30	Lack of clinical trial focus
Edwards JG, Dinan TG, Waller DG, Greentree SG. Double-blind comparative study of the antidepressant, unwanted and cardiac effects of minaprine and amitriptyline. Br J Clin Pharmacol 1996;42(4):491-8	Incomplete relevant data reported
Emslie G, Kratochvil C, Vitiello B, Silva S, Mayes T, McNulty S, Weller E, Waslick B, Casat C, Walkup J, Pathak S, Rohde P, Posner K, March J;Columbia Suicidality Classification Group; TADS Team. Treatment for Adolescents with Depression Study (TADS): safety results. J Am Acad Child Adolesc Psychiatry 2006;45(12):1440-55	Not possible to compare data between methods
Fisher S, Bryant SG. Postmarketing surveillance: accuracy of patient drug attributionjudgments. Clin Pharmacol Ther 1990;48(1):102-7	Lack of clinical trial focus
Gelenberg AJ, Dunner DL, Rothschild AJ, Pedersen R, Dorries KM, Ninan PT. Sexual functioning in patients with recurrent major depressive disorder enrolled in the PREVENT study. J Nerv Ment Dis 2013;201(4):266-73	Not possible to compare data between methods
Glaser EM, Whittow GC. Experimental errors in clinical trials. Clin Sci (Lond) 1954;13(2):199-210	Not possible to compare data between methods
Greenblatt M. Controls in clinical research. Clin Pharmacol Ther 1964;5(6):864-70	Incomplete relevant data reported
Hakobyan L, Haaijer-Ruskamp FM, de Zeeuw D, Dobre D, Denig P. A review of methods used in assessing non-serious adverse drug events in observational studies among type 2 diabetes mellitus patients. Health Qual Life Outcomes 2011;9:83	Lack of clinical trial focus
Hanesse B, Legras B, Royer RJ, Guillemin F, Briancon S. Adverse drug reactions: comparison of two report methods. Pharmacoepidemiology and drug safety 1994;3:223-	Lack of clinical trial focus
Homsi J, Walsh D, Rivera N, Rybicki LA, Nelson KA, Legrand SB, Davis M, Naughton M, Gvozdjan D, Pham H. Symptom evaluation in palliative medicine: patient report vs systematicassessment. Support Care Cancer 2006;14(5):444-53	Lack of clinical trial focus
Iversen P, Karup C, van der Meulen E, Tankó LB, Huhtaniemi I. Hot flushes in prostatic cancer patients during androgen-deprivation therapy with monthly dose ofdegarelix or leuprolide. Prostate cancer and prostatic diseases 2011;14(2):184-9	Incomplete relevant data reported
Jarernsiripornkul N, Kakaew W, Loalukkana W, Krska J. Adverse drug reaction monitoring: comparing doctor and patient reporting for new drugs. Pharmacoepidemiol Drug Saf 2009;18(3):240-5	Lack of clinical trial focus
Jonsson T, Christrup LL, Højsted J, Villesen HH, Albjerg TH, Ravn-Nielsen LV, Sjøgren P. Symptoms and side effects in chronic non-cancer pain: patient report vs. systematic assessment. Acta Anaesthesiol Scand 2011;55(1):69-74	Lack of clinical trial focus
Lambert TJ, Cock N, Alcock SJ, Kelly DL, Conley RR. Measurement of antipsychotic-induced side effects: support for the validity of a self-report (LUNSERS) versus structured interview (UKU) approach to measurement. Hum Psychopharmacol 2003;18(5):405-11	Included objective measure(s)
Love RR, Leventhal H, Easterling DV, Nerenz DR. Side effects and emotional distress during cancer chemotherapy. Cancer 1989;63(3):604-12	Lack of clinical trial focus
Makarananda K, Sriwatanakul K, Pothisiri P, Eamrungroj S, Charoonroj P, Pongprayoon U. Prospective study of adverse drug reactions of antihypertensive drugs in Thai outpatients. J Med Assoc Thai	Lack of clinical trial focus

1985;68(5):229-36	
Martys CR. Monitoring adverse reactions to antibiotics in general practice.. J Epidemiol Community Health 1982;36(3):224-7	Lack of clinical trial focus
Mei PA, Montenegro MA, Guerreiro MM, Guerreiro CA. Pharmacovigilance in epileptic patients using antiepileptic drugs.. Arq Neuropsiquiatr 2006;64(2A):198-201	Lack of clinical trial focus
Möller HJ, Glaser K, Leverkus F, Göbel C. Double-blind, multicenter comparative study of sertraline versus amitriptyline in outpatients with majordepression. Pharmacopsychiatry 2000;33(6):206-12	Not possible to compare data between methods
Olsen H, Klemetsrud T, Stokke HP, Tretli S, Westheim A. Adverse drug reactions in current antihypertensive therapy: a general practice survey of 2586 patients in Norway. Blood Press 1999;8(2):94-101	Lack of clinical trial focus
Pandina GJ, Bossie CA, Zhu Y, Gharabawi GM. Evaluating movement disorders in pediatric patients receiving risperidone: a comparison of spontaneous reports and research criteria for TD. Child Adolesc Psychiatry Ment Health 2007;1(1):3	Included objective measure(s)
Rynn MA, Walkup JT, Compton SN, Sakolsky DJ, Sherrill JT, Shen S, Kendall PC, McCracken J, Albano AM, Piacentini J, Riddle MA, Keeton C, Waslick B, Chrisman A, Iyengar S, March JS, Birmaher B. Child/Adolescent anxiety multimodal study: evaluating safety. J Am Acad Child Adolesc Psychiatry 2015;54(3):180-90	Not possible to compare data between methods
Sheikh SI, Nestorov I, Russell H, O'Gorman J, Huang R, Milne GL, Scannevin RH, Novas M, Dawson KT. Tolerability and pharmacokinetics of delayed-release dimethyl fumarate administered with and without aspirin in healthy volunteers. Clin Ther 2013;35(10):1582-94	Not possible to compare data between methods
Thomsen HS. Frequency of acute adverse events to a non-ionic low-osmolar contrast medium: the effect of verbal interview. Pharmacol Toxicol 1997;80(2):108-10	Lack of clinical trial focus
Tran PV, Dellva MA, Tollefson GD, Beasley CM Jr, Potvin JH, Kiesler GM. Extrapyramidal symptoms and tolerability of olanzapine versus haloperidol in the acute treatment of schizophrenia. J Clin Psychiatry 1997;58(5):205-11	Included objective measure(s)
Trindade E, Menon D, Topfer LA, Coloma C. Adverse effects associated with selective serotonin reuptake inhibitors and tricyclic antidepressants: a meta-analysis. CMAJ 1998;159(10):1245-52	Comparison between trials, not within trials
Van Haecht CH, Vander Stichele R, Bogaert MG. Package inserts for antihypertensive drugs: use by the patients and impact on adverse drug reactions. Eur J Clin Pharmacol 1990;39(6):551-4	Lack of clinical trial focus
Waddell L, Taylor M. A new self-rating scale for detecting atypical or second-generation antipsychotic side effects.. J Psychopharmacol 2008;22(3):238-43	Lack of clinical trial focus
Yusufi B, Mukherjee S, Flanagan R, Paton C, Dunn G, Page E, Barnes TR. Prevalence and nature of side effects during clozapinemaintenance treatment and the relationship with clozapine dose and plasma concentration. Int Clin Psychopharmacol 2007;22(4):238-43	Included objective measure(s)

Supplementary Table S3.3: Characteristics of studies awaiting classification (Chapter 3)

Aspirin Myocardial Infarction Study research group. A randomized, controlled trial of aspirin in persons recovered from myocardial infarction. JAMA 1980;243:661-9	
Methods	Within-participant comparison of data elicited by 2 methods. Adults who had had a myocardial infarction enrolled in a randomised, double-blind, placebo-controlled trial of aspirin.
Data	AEs (haematemesis, tarry stools, bloody stools).
Comparisons	Open question versus specific enquiry.
Outcomes	Proportion of participants reporting AEs by method by trial group.
Notes	This comparison was referenced by LM Friedman in Fundamentals of Clinical Trials (Springer), however none of the published papers found so far for this clinical trial reported the comparison data.
Mothapo KM, Schellekens A, Van Crevel R, Keuter M, Grintjes-Huisman K, Koopmans P, Der Ven VA.. Improvement of Depression and Anxiety after Discontinuation of Long-Term Efavirenz Treatment. CNS Neurol Disord Drug Targets 2015	
Methods	Within-participant comparison of data elicited by 3 methods. HIV-infected participants on long-term efavirenz in an observational clinical trial.
Data	Neuropsychiatric symptoms
Comparisons	The depression-anxiety-stress-scale (DASS), the symptom-checklist (SCL-90) and the outcome-questionnaire (OQ-45).
Outcomes	Not clear
Notes	Awaiting access to full text.

Supplementary Table S3.4: Detail of included studies' questioning methods (Chapter 3)

	Description	Application	Development
Open questions			
Allen 2013	Questions about health and medicine-use without reference to particular conditions, body systems or treatments.	Verbal. Answers probed according to common clinical practice in eliciting a medical and treatment history.	No details
Avery 1967	"Have you noticed any change in bodily function or had any physical complaints in the past week?".	Verbal.	No details
Barber 1995	Participants instructed to call if they experienced AE and asked at visits if they had experienced AE.	Passive and verbal.	No details
Barrowman 1970	Asked to describe any unusual sensation.	Verbal.	
Bent 2006	“Did you have any significant medical problem since the last study visit?”). If “yes”, asked to identify.	Self-administered, recorded by study assistant on checklist.	No details
Bent 2006	“Since the last study visit, have you limited your usual daily activities for more than 1 day because of a medical problem?”). If “yes”, asked to identify.	Self-administered, recorded by study assistant on checklist.	No details
Borghi 1984	No detail.	No details	No detail
Brent 2009	Spontaneous reports.	No details	No details
Downing 1970	"How are you feeling?". If no reference to drug-related symptomology: "How else are you feeling?" then "How does the drug make you feel?".	No details	No details

Greenhill 2004	"Has ____ had any physical or health problems since….? I'm talking about something that started to become a problem during this time or an old problem that got much worse.", "Have there been activities that ____ didn't do as often or that he/she didn't do at all because of not feeling well since _____?", Since ____, has ____ said that his/her body feels funny... or that he/she has any aches or pains... or that some part of him/her hurts or doesn't feel well?"	Verbal.	Constructed from an existing instrument (SAFTEE) as a scripted interview, with instructions, demographic queries and a glossary of preferred AE terms.
Hermans 1994	"How have you felt since your last visit?"	Verbal.	No details
Huskisson 1974	"Have you noticed any new symptoms which might be related to the treatment?"	Verbal.	No details
Jacobson 1987	"Have you had any physical or health problems during the past week [or specified assessment interval]? Have you noticed any changes in your physical appearance during the past week [or specified assessment interval]? Have you cut down on the things you usually do because of not feeling well physically during the past week [or specified assessment interval]?"	Verbal structured interview.	SAFTEE developed by the National Institute of Mental Health. This study part of validation exercise.
Kruft 2007	"How are you doing since your last visit?".	No details	No details
Landén 2005	Non-leading question such as: "Have you felt different in any way since you started the new treatment?"	No details	No details
Monteiro 1987	Spontaneous reports.	No details	No details
Nicholls 1980	Spontaneous reports.	No details	No details
O'Connell 2007	Asked to list any side effects or changes experienced.	No details	No details

Os 1994	Spontaneous reports.	No details	No details
Perez-Lloret 2012	"Have you noticed any unpleasant effects of your medications during the previous week?"	No details	No details
Rabkin 1992	"Have you had any physical or health problems during the past week [or specified assessment interval]? Have you noticed any changes in your physical appearance during the past week [or specified assessment interval]? Have you cut down on the things you usually do because of not feeling well physically during the past week [or specified assessment interval]?	Verbal structured interview.	SAFTEE developed by the National Institute of Mental Health. Additional 11 items research team's own choice. Severity subjectively graded and action taken noted.
Reilly 1992	Asked about any health-related symptoms or problems.	No details	No details
Rosenthal 1996	Physician interview concerning general status of symptoms, development of signs (e.g. rash, swelling, bruises).	No details	No details
Török 1984	Spontaneous reports.	No details	No details
Wallander 1991	"Have you had any health problems since we first met?"	No details	No details
Wallin 1981	"Have you had any troubles from the drug?"	No details	No details
Wernicke 2005	Asked to report experiences in own words.	No details	No details
Yeo 1991	"Have the tablets upset you in any way?"	No details	No details
Blank forms			
Borghi 1984	Blank card to report signs and symptoms experienced.	Self-completed.	No details
Ciccolunghi 1975	Open-ended questionnaire with 3 entry lines.	No details	No details
Ciccolunghi 1975	Open-ended questionnaire with 10 entry lines.	No details	No details

Spilker 1987	Questionnaire with 15 blank spaces to complete about demographics, tobacco and alcohol use, symptoms experienced, treatments used.	Self-completed.	No details
Sheftell 2004	Asked if they had AEs when using drug. If yes, asked to list and grade severity.	No details	No details
Checklists			
Allen 2013	Potential health issues and medicines by10 body systems, 28 symptoms, 17 medicines in total for both trials together.	Verbal enquiry. Answers probed according to common clinical practice in eliciting a medical and treatment history.	A majority of fields were common between the 2 trials, although they could not be harmonised fully.
Avery 1967	Possible drug side effects.	No details	National Institute of Mental Health study checklist (NB assumed not all were asked).
Barber 1995	Common side effects.	Interviewer-administered.	Validated questionnaire for capturing frequency and bother of ocular and other local effects, effects on visual function of topical therapy for lowering intraocular pressure. Also, extent to which side effects and associated limitations in routine living activities interfere with health-related quality of life, medication compliance, and patient satisfaction with the medication.
Barrowman 1970	Symptoms.	Verbal enquiry	Symptoms known or suspected to occur after pentagastrin, and some control items (headache, dryness of mouth, increased salivation).
Bent 2006	“Since the last visit, have you experienced any of the following?”: 53 symptoms, grouped by anatomical region	Self-administered.	Developed after a unpublished review of checklists used in earlier trials at same institution.
Borghi 1984	49 items requiring yes/no.	Self-completed (sequence changed each visit)	Pharmacological unwanted effects linked to the most common antihypertensive drugs, mainly ß-blockers and diuretics.
Ciccolunghi 1975	38 items	No details.	No details

de Vries 2013	Adverse drug events (ADEs) categorised in 16 body categories.	Email invite for Internet-based self-administration.	Content validation of common ADEs drafted in layman terms with reference to CTCAE v4.0, existing symptom and ADE checklists, then coded, categorised into body categories using MedDRA® classifications. Open-ended option for "other", questions relating to duration, frequency, seriousness and causality based on literature.
de Vries 2013	Adverse drug events not categorised in body categories.	Email invite for Internet-based self-administration.	Content validation of common ADEs drafted in layman terms with reference to CTCAE v4.0, existing symptom and ADE checklists, then coded, categorised into body categories using MedDRA® classifications. Open-ended option for "other", questions relating to duration, frequency, seriousness and causality based on literature.
de Vries 2014	Adverse drug events (symptoms in lay terms).	Email invite for Internet-based self-administration.	Content validation of common ADEs drafted in layman terms with reference to CTCAE v4.0, existing symptom and ADE checklists, then coded, categorised into body categories using MedDRA® classifications. Open-ended option for "other", questions relating to duration, frequency, seriousness and causality based on literature.
Downing 1970	28-item questionnaire. If AE reported, patient asked to estimate intensity, discomfort and whether symptom was felt due to study medicine.	Structured interview using a coding sheet.	23 common medication effects, and 5 highly unlikely to be related captured as miscellaneous.
Greenhill 2004	Drug-specific inquiry - 18 questions	Verbal enquiry	Constructed from an existing instrument (SAFTEE) as a scripted interview, with instructions, demographic queries and a glossary of preferred AE terms. Clinically important AEs for various medicines.
Greenhill 2004	Body system review - 24 questions	Verbal enquiry	Constructed from an existing instrument (SAFTEE) as a scripted interview, with instructions, demographic queries and a glossary of preferred AE terms. Clinically important AEs for various medicines.

Hermans 1994	Symptoms.	Self-completed written questionnaire	Anticipated side-effects of dihydropyridine calcium antagonists (swollen ankles, headache, flushing, palpitations, dizziness or nausea).
Huskisson 1974	21 possible side effects	No details	Tinnitus, deafness, gastrointestinal complaints and others with no obvious relevance and others with no obvious relevance.
Jacobson 1987	Review of 23 body systems.	Verbal structured interview.	SAFTEE developed by the National Institute of Mental Health. This study part of validation exercise.
Kruft 2007	Ophthalmic symptoms.	No details	No details
Lundberg 1980	Somesthetic Inventory: 54 body feelings (NB 9 fillers assessed through a visual analogue scale.)	Self-completed - participants asked to 'attune to inner stimuli', close eyes and determine how body felt.	Compilation and organization of data on side-effects of antihistamines
Monteiro 1987	Self-rated physical symptom questionnaire (including sexual function items).	No details	No details
Nicholls 1980	Side-effects	Self-administered	No details
O'Connell 2007	12 specific side effects		AEs commonly attributed to OCs, including headache, nausea, acne, abdominal pain, back pain, vomiting, breast tenderness, breast enlargement, mood swings, weight gain, premenstrual syndrome and irregular bleeding
Os 1994	34 symptomatic side-effects.	Direct questioning to be answered 'yes' or 'no'.	Part of the ASPECT Scale - a tool for evaluation of 34 commonly experienced symptomatic side-effects of cardiovascular drugs.
Perez-Lloret 2012	Pre-defined list of AEs.	Verbal structured enquiry.	pre-defined list of most common ADRs to various anti-PD drugs from a literature search critically reviewed by a group of PD and pharmacovigilance specialist for consensus: general, GI, urinary, neuropsychiatric, dermatologic
Rabkin 1992	23 body systems plus additional 11 items to represent side effects of MAOIs.	Verbal structured interview.	SAFTEE developed by the National Institute of Mental Health. Additional 11 items research team's own choice. Severity subjectively graded and action taken noted.

Reilly 1992	4 symptoms.		4 symptoms associated with hypertension and anti-hypertensive therapy (previously used)
Rosenthal 1996	Signs and symptoms.	Self-administered written questionnaire	Formulated to elicit information concerning the appearance of signs and symptoms that could be related to ACE inhibitors or a beta-blocker, and any other symptoms that may reflect well being.
Sheftell 2004	49 possible AEs.		Mostly known Triptan side effects and some confounders (side effects not expected to be related with triptans).
Spilker 1987	25 symptoms.	Self-completed	No details
Török 1984	Side-effects.	No details	35 anticipated and other side-effects
Wallander 1991	"Have you had any of the following symptoms in the past month?"	No details	Used in previous population studies, includes depression, tension, head, heart, lung, metabolism, musculoskeletal system, GI and urinary tracts. Patients place in an envelope and advised the physician would not have access
Wallin 1981	"Have you noticed any of the following reactions: diarrhoea, nausea, vomiting, other gastrointestinal disturbances, skin eruptions, or other troubles?"	No details	No details
Wernicke 2005	Side Effects Checklist (child trial), BBAEQ-M (child/adolescent trial), AMDP-5 (adult trial)	No details	Side Effects Checklist is based on the Subjective Treatment Emergent Symptoms Scale (US National Institute of Mental health) - 30 items including general symptoms such as trouble sleeping, diarrhoea, headache, trouble eating. BBAEQ-M: 24 items rated 0-9, AMDP-5: 47 items rated 0-3.
Rating scales			
Brent 2009	Brief Suicide Severity Rating Scale: rating of suicidal ideation 0 to 5 and rating of suicidal behavior 0 to 5 using Columbia Classification Algorithm of Suicide Assessment.	No details	Published validated instruments.
Kruft 2007	Visual analogue scales (VAS).	No details.	Some VAS were validated instruments, details

			unknown.
Landén 2005	UKU side-effect rating scale (none, mild, moderate, severe) for 3 symptoms of sexual dysfunction.	No details	UKU is a validated instrument.
Lundberg 1980	Side Effects Report of 24 terms assessed with VAS.	Self-completed - definitions provided to participants.	Arbitrarily selected terms.
Os 1994	VAS.	Completed by participant and spouse independently	No details
Wallander 1991	41 item VAS.	No details	Validated instrument, highly correlated items in 6 domains, rest single items
Yeo 1991	VAS (cough).	No details	No details
Diary			
de Vries 2014	Open-ended question asking for symptoms experienced and closed-ended question about attribution to any drug taken.	Paper-based.	
In-depth interview			
Allen 2013	Prompted narrative of participant's trial experience, reflection on previous ill health and medicines used and photographs of typical over-the-counter and traditional medicines available to the study populations.	Verbal interview.	No details
Monteiro 1987	Stuctured interview.	Verbal interview.	No details.

Supplementary Table S3.5a: Risk of bias of included studies for within-participant comparisons (Chapter 3)

	Selection bias				Performance and detection bias		Attrition bias		Reporting bias					
Author	**Random sequence generation**		**Allocation concealment**		**Blinding**		**Incomplete outcome data assessed**		**No selective reporting**		**Explicit application**		**Other**	
Allen 2013	N/A		High risk	Open by nature of study	High risk	Participants/assessors unlikely to be blinded due to the nature of study. 'Possible 'priming' by earlier method, however cumulative data was part of the study	Low risk	Appears all participants completed all methods	Low risk	All data appear presented	Unclear	Possible variations in phraseology between participants.	N/A	
Barber 1995	N/A		High risk	Open by nature of study	High risk	Participants/assessors unlikely to be blinded due to the nature of study. 'Possible 'priming' by earlier method	Low risk	A few drop outs, unlikely related to methods	Low risk	All data appear presented	Unclear	Possible variations in phraseology between participants.	Unclear	Query if methods applied in same order for all participants
Barrowman 1970	N/A		High risk	Open by nature of study	High risk	Participants/assessors unlikely to be blinded due to the nature of study. 'Possible 'priming' by earlier method	Low risk	Appears all participants completed all methods	Low risk	All data appear presented	Low risk	Explicit instructions for staff to use	N/A	
De Vries 2013	Low risk	Random sequence by two groups	High risk	Open by nature of study	High risk	Participants/assessors unlikely to be blinded due to the nature of study. 'Possible 'priming' by earlier method	Low risk	A few drop outs, unlikely related to methods	Low risk	All data appear presented	Low risk	Assume explicit instructions for self-administration.	N/A	

De Vries 2014	N/A		High risk	Open by nature of study	High risk	Participants/assessors unlikely to be blinded due to the nature of study. 'Possible 'priming' by earlier method	Unclear	Not clear whether attrition was related to method	Low risk	All data appear presented	Low risk	Assume explicit instructions for self-administration.	N/A	
Downing 1970	N/A		High risk	Open by nature of study	High risk	Participants/assessors unlikely to be blinded due to the nature of study. 'Possible 'priming' by earlier method	Low risk	Appears all participants completed all method	Unclear	Only presented data that patients related to medication	Low risk	Explicit instructions for staff to use	N/A	
Greenhill 2004	N/A		High risk	Open by nature of study	High risk	Participants/assessors unlikely to be blinded due to the nature of study. 'Priming' by earlier method, however cumulative data was part of the study	Low risk	Appears all participants completed all method	Low risk	All data appear presented	Low risk	Explicit instructions for staff to use	N/A	
Hermans 1994	N/A		High risk	Open by nature of study	High risk	Participants/assessors unlikely to be blinded due to the nature of study. Possible 'priming' by earlier method	Low risk	A few drop outs, unlikely related to methods	Low risk	All data appear presented	Low risk	Assume explicit instructions, particularly for self-administration	N/A	
Jacobson 1987	Low risk	Random sequence order of questions received by two groups	High risk	Open by nature of study	High risk	Participants/assessors unlikely to be blinded due to the nature of study. Possible 'priming' by earlier method	Unclear	No information provided	Unclear	Some overlap between data presented	Low risk	Explicit instructions for staff to use	N/A	
Kruft 2007	N/A		High risk	Open by nature of study	High risk	Participants/assessors unlikely to be blinded due to the nature of study. Possible 'priming' by earlier method	Unclear	No information provided	Unclear	Difficult to ascertain as short report	Unclear	Not clear from information provided.	Unclear	Meta-analysis

Landen 2005	N/A		High risk	Open by nature of study	High risk	Participants/assessors unlikely to be blinded due to the nature of study. Possible 'priming' by earlier method	Low risk	Appears all participants completed all methods	Unclear	Summary data presented	Low risk	Assume explicit instructions, particularly for self-administration.	N/A	
Lundberg 1980	N/A		High risk	Open by nature of study	High risk	Participants/assessors unlikely to be blinded due to the nature of study. Possible 'priming' by earlier method	Low risk	Appears all participants completed all methods	Unclear	Summary data presented	Low risk	Explicit instructions for staff to use.	N/A	
Montiero 1987	N/A		High risk	Open by nature of study	High risk	Participants/assessors unlikely to be blinded due to the nature of study. Possible 'priming' by earlier method	Low risk	A few drop outs, unlikely related to methods	Unclear	Summary data presented	Unclear	Not clear from information provided.	N/A	
Nicholls 1980	N/A		High risk	Open by nature of study	High risk	Participants/assessors unlikely to be blinded due to the nature of study. Possible 'priming' by earlier method	Unclear	Not clear why different numbers of participants for 2 methods	Low risk	All data appear presented	Unclear	Not clear from information provided.	N/A	
O'Connell 2007	N/A		High risk	Open by nature of study	High risk	Participants/assessors unlikely to be blinded due to the nature of study. Possible 'priming' by earlier method	Low risk	Appears that all participants completed all methods	Low risk	All data appear presented	Unclear	Not clear from information provided.	N/A	
Os 1994	N/A		High risk	Open by nature of study	High risk	Participants/assessors unlikely to be blinded due to the nature of study. Possible 'priming' by earlier method	Low risk	Some drop outs but unlikely related to method	Low risk	All data appear presented	Unclear	Not clear from information provided.	Unclear	Unclear order

Perez-Lloret 2012	N/A		High risk	Open by nature of study	High risk	Participants/assessors unlikely to be blinded due to the nature of study. Possible 'priming' by earlier method	Low risk	Appears that all participants completed all methods	Low risk	All data appear presented	Low risk	Explicit instructions for staff to use	N/A	
Rabkin 1992	N/A		High risk	Open by nature of study	High risk	Participants/assessors unlikely to be blinded due to the nature of study. Possible 'priming' by earlier method	Low risk	Appears that all participants completed all methods	Low risk	All data appear presented	Low risk	Explicit instructions for staff to use	N/A	
Reilly 1992	N/A		High risk	Open by nature of study	High risk	Participants/assessors unlikely to be blinded due to the nature of study. Possible 'priming' by earlier method	Low risk	Some drop outs but unlikely related to method	Low risk	All data appear presented	Unclear	Not clear from information provided	N/A	
Rosenthal 1996	N/A		High risk	Open by nature of study	High risk	Participants/assessors unlikely to be blinded due to the nature of study. Possible 'priming' by earlier method	Unclear	Not clear why less participants for method 2, could be related to matching of symptoms between methods	High risk	Not all data presented - those spontaneously-elicited were only presented if they matched questions on the questionnaire	Unclear	Not clear from information provided	N/A	
Sheftell 2004	N/A		High risk	Open by nature of study	High risk	Participants/assessors unlikely to be blinded due to the nature of study. Possible 'priming' by earlier method	Low risk	Appears that all participants completed all methods	Low risk	All data appear presented	Low risk	Assume explicit instructions for self-administration.	N/A	

Wallander 1991	N/A		High risk	Open by nature of study	High risk	Participants/assessors unlikely to be blinded due to the nature of study. Possible 'priming' by earlier method, especially as participants took forms for both questioning methods home to complete	High risk	Significant number of drop outs (45/236), potentially related to method (participants took forms for both questioning methods home to complete)	Low risk	All data appear presented	High risk	Although can assume explicit instructions for self-administration of methods 2 and 3, they were completed at home so it was outside of the control of the clinic as to completion, including order.	N/A	
Wallin 1981	N/A		High risk	Open by nature of study	High risk	Participants/assessors unlikely to be blinded due to the nature of study. Possible 'priming' by earlier method	Unclear	No information provided	Unclear	Not clear from information provided	Unclear	May have been variations in phraseology between participants in method 2.	N/A	
Wernicke 2006	N/A		High risk	Open by nature of study	High risk	Participants/assessors unlikely to be blinded due to the nature of study. Possible 'priming' by earlier method	Unclear	No information provided	High risk	Not all data presented. AEs reflecting same symptom in spontaneous and solicited methods were selected for the comparison	Unclear	May have been variations in phraseology between participants in method 1.	Unclear	Meta-analysis
Yeo 1991	N/A		High risk	Open by nature of study	High risk	Participants/assessors unlikely to be blinded due to the nature of study. Possible 'priming' by earlier method	Low risk	Some drop outs (3 more in method 2) but unlikely related to method	Low risk	All data appear presented	Low risk	Standard O and assume explicit instructions for self-administration.	N/A	

Supplementary Table S3.5b: Risk of bias of included studies for between-participant comparisons (Chapter 3)

	Selection bias				Performance and detection bias		Attrition bias		Reporting bias					
Author	**Random sequence generation**		**Allocation concealment**		**Blinding**		**Incomplete outcome data assessed**		**No selective reporting**		**Explicit application**		**Other**	
Avery 1967	Unclear	No information other than "arbitrarily assigned"	Unclear	No information provided	High risk	No information provided, however participants and assessors unlikely to be blinded due to the nature of study. The group who were exposed to both questioning methods may have been 'primed' by the first method	Low risk	Drops outs were indicated but appear similar between groups	Unclear	Summary data presented	Low risk	Explicit phraseology for O. CL had specific items but not all is presented in laymans language so may have had inconsistent phraseology	N/A	
Bent 2006	Low risk	Computer-generated prior to study. Baseline characteristics were similar between groups	Low risk	Study personnel were blinded to the allocation	High risk	Although participants were not informed of their group and analysts were blinded, study staff were aware of groups and were involved in completing some data	Low risk	All participants completed the study and outcome assessment	Low risk	All data appear presented	Low risk	Assume explicit instructions for self-administration	N/A	
Borghi 1984	Unclear	No information provided	Unclear	No information provided	Low risk	Little information provided, however the investigator was neither informed of the results of the self-reporting, nor did they help participants fill in the forms	Low risk	Although group allocation of drop outs unclear, there were only a few	Low risk	All data appear presented	Unclear	Not clear from information provided	N/A	

Brent 2009	High risk	Non-random allocation by nature of study	High risk	Open allocation by enrolment period and nature of study	High risk	Participants and assessors unlikely to be blinded due to the nature of study	Unclear	No information provided	Unclear	Brief summary data presented	Unclear	Little information provided, although method 2 involved validated scales so assumed had explicit instuctions on application	Unclear	Unclear whether method 1 was used in both groups
Chiccolunghi 1975	Low risk	Pre-determined randomisation list. Baseline characteristics were similar between groups	Unclear	No information provided	High risk	Participants and assessors unlikely to be blinded due to the nature of study	High risk	Significant missing data, potentially related to the method as forms were distibuted in the internal mail and it was left to staff to decide whether to complete and return them, although simliar numbers were returned for each elicitation type (57 vs 44%)	Low risk	All data appear presented	Low risk	Assume explicit instructions for self-administration	Unclear	Large amount of non-responders to invitation which could be related to methods
Huskisson 1974	High risk	Non-random allocation by study centre	High risk	Open by nature of study	High risk	Participants and assessors unlikely to be blinded due to the nature of study, although groups were at different sites there may still be room for biased assessments based on the method of questioning, and different staff may elicit /record AEs differently	Unclear	No information provided	Unclear	Raw data transformed by scoring and some data grouped	Unclear	Not clear from information provided	Unclear	Unclear whether method 1 used in both groups

Spilker 1987	Low risk	Table of random numbers	Unclear	No information provided	High risk	Participants and assessors unlikely to be blinded due to the nature of study	Low risk	Although group allocation of drop outs unclear, there were only a few	Low risk	All data appear presented	Low risk	Assume explicit instructions for self-administration	Unclear	How many were invited not stated
Torok 1984	High risk	Non-random allocation by nature of study	High risk	Open by nature of study	High risk	Participants and assessors unlikely to be blinded due to the nature of study, although groups were at different sites/in different studies there may still be room for biased assessments based on the method of questioning	Unclear	No information provided	Unclear	Relevant AE data only presented for participants taking chloranolol	Unclear	Not clear from information provided	N/A	

Supplementary Table S3.6a: Effects of methods on the number of AEs reported for between-participant comparisons (Chapter 3)

Study	Therapy area	Endpoint	Follow-up	AEs elicited												Effect (number of AEs)		
				Number participants				Number of AEs (total)				Number (%) participants with ≥1 AE				Overall		By drug arm
				O	O	CL	R	O	O	CL	R	O	O	CL	R	Proportion with ≥ 1 AE†	Description (including of other effect measure if different)	
Avery 1967	Psychiatry	Any AE	5 weeks	11	N/A	12	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	Statistically significant higher mean number of AEs at each visit by CL. See paper for details	Statistically significant higher mean number of AEs at 5 of 6 study visits by CL in the active drug arm. See paper for details.
Huskisson 1974**	Rheumatology	Any AE	24 weeks	30	N/A	30	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	Unclear scoring; appears higher total score for AEs by CL than O (540 vs 409). AEs listed on CL were more frequently reported by CL than when not listed. AEs not listed on CL more frequently reported by O. See paper for details	Fenoprofen auditory, gastro-intestinal (GI) I and all other ('irrelevant') AE scores 2-3 x more frequently reported by CL than O, aspirin AE scores ranged from no difference (auditory), approx.. 1.5 x less (GI) and 0.60 x more ('irrelevant') using O compared to CL. See paper for details.
Torok 1984***	Cardiology	Any AE	Various	600	N/A	537, 929	N/A	7	N/A	75, 365	N/A	N/A	N/A	N/A	N/A	N/A	Lower number of AEs per 100 participants by O compared to CL	N/A

Bent 2006	Prostatic hyperplasia	Any AE	< 1 - 24 months	70	70	74	N/A	11	14	238	N/A	10(14)	9(13)	57(77)	N/A	OvsO 0.14 (-0.10;0.12) P=0.805, OvsCL -0.63 (-0.75;-0.50) P=0.000, OvsCL -0.64 (-0.77;-0.52) P=0.000	Group assigned to CL reported significantly greater number of AEs than either O. No difference between O methods	N/A
Borghi 1984*	Cardiology	Any AE	16 weeks	106	N/A	117	N/A	496	N/A	1556	N/A	Oxprenolol 48 (45), chlorthalidone 60 (57)	N/A	Oxprenolol 76 (65), chlorthalidone 81 (69)	N/A	Only available by drug arm	N/A	% ≥ 1 AE for Oxprenolol 0.20 (0.07;0.33) P=0.0031, Chlorthalidone 0.13 (0.00;0.25) P=0.05
Spilker 1987	No indication	Any AE	1 occasion	132	N/A	166	N/A	229	N/A	581	N/A	106 (80)	N/A	154 (93)	N/A	-0.12 (-0.20;-0.05) P=0.013	Group assigned to CL reported significantly greater number of AEs than O	N/A
Ciccolunghi 1975	No indication	Any AE	1 occasion	144	129	143	N/A	88	67	720	N/A	59 (41)	46 (36)	127 (89)	N/A	OvsO 0.05 (-0.06;0.16) P=0.3676, OvsCL -0.48 (-0.57;-0.38) P=0.000, OvsCL -0.53 (-0.63;-0.43) P=0.000	Group assigned to CL reported significantly greater number of AEs than either O. No difference between O methods	% ≥ 1 AE for non-med: OvsO 0.19 (-0.09;0.13) P=0.7331, OvsCL -0.66 (-0.77; -0.54) P=0.0000, OvsCL -0.68 (-0.79;-0.56) P=0.0000. Med: OvsO 0.05 (-0.12;0.22) P=0.5816, OvsCL -0.26 (-0.38;-0.14) P= 0.001, Ovs CL -0.31 (-0.44;-0.17) P=0.0000
Brent 2009**	Psychiatry	Self-harm	12 weeks															

Suicidal				181	N/A	N/A	153	N/A	N/A	N/A	N/A	16(8.8)	N/A	N/A	32(20.9)	-0.12 (-0.22;0.09) P=0.0017	Group assigned to R reported significantly greater number of suicidal-related AEs than O. NB o completed suicides	N/A
Non-suicidal				181	N/A	N/A	153	N/A	N/A	N/A	N/A	4(2.2)	N/A	N/A	27(17.6)	-0.15(-0.22;-0.90) P=0.0000	Group assigned to R reported significantly greater number of non-suicidal AEs than O	N/A
Suicide attempts				181	N/A	N/A	153	N/A	N/A	N/A	N/A	7(3.9)	N/A	N/A	10(6.5)	-0.03 (-0.07;0.02) P=0.2689	No difference between R and O for reporting of suicide attempts	N/A

O: open question/spontaneous, O(B): blank page, CL: checklist, R: rating scale

* All participants asked AEs by an open question (O), but as this process possibly involved 'filtering' of reports by the doctor the data were not included in the review.

** Huskisson used a composite measure of frequency and severity

*** A selection of the total AEs are presented in the review as those objectively measured were excluded

† 2-sample test of proportions

†† 2-sample t test

Supplementary Table S3.6b: Effects of methods on the number of AEs reported for within-participant comparisons (Chapter 3)

Study	Therapy area	Endpoint	Follow-up	AEs elicited							Effect (number of AEs)								
				Number of participants	Number of AEs (total)						Number (%) of participants with ≥1 AE						Overall		By drug arm
					O	O	CL	CL	R	INT	O	O	CL	CL	R	INT	Test of proportions with ≥ 1 AE	Description (including of other effect measure if different)	
Barber 1995	Ophthalmology	Any AE	4 weeks	92	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	Average frequency of domain scores [number (%) mean (SD)] for those not reporting AEs by O but indicating AE by CL: occular symptoms 41 (89.1), 1.18 (0.91); taste 3 (6.4), 2.5 (2.18); vision difficulties 33 (70.2), 2.8 (1.84); accommodation difficulties 20 (42.6), 3.68 (2.27); browache 12 (25.5), 2.75 (1.86). Scores range from participants experiencing AE rarely to usually. Average domain scores increased as AEs reported by O and therapy discontinued - participants reporting AEs to O reported more AEs by CL	N/A
Barrowman 1970	GI	Any AE	1 occasion	24	31	N/A	57	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	CL elicited 1.8 x the number of AEs than O	Mean number of AEs for pentagastrin: O 2.1 CL 3.2 (range 1-5). For placebo: O 0

																			CL 1 (range 0-3).
Downing 1970	Psychiatry	Any AE	4 weeks	123	N/A	N/A	N/A	N/A	N/A	N/A	45	N/A	65	N/A	N/A	N/A	OR 3.22 (1.49; 7.74) P=0.0017†	While methods agreed for 85 (69%) of participants, CL elicited a significantly greater number of AEs than O	More AEs reported with amplodipine versus isradipine when either O or CL used No drug-placebo difference found when O used, however for CL a statistically significant difference found in proportion of participants with ≥ 1 AE (χ^2 5.76 P < 0.025). In those AEs deemed medication-related, CL resulted in higher frequency of AEs in active arm and produced a larger drug-placebo difference in frequency of AEs
Hermans 1994	Cardiology	Any AE	6 weeks	205	W2: 11 W4: 49	N/A	W2: 58 W4: 105	N/A	N/A	N/A	W2: 21 (10) W8: 53 (26)	N/A	W2: 39 (19) W8: 61 (31)	N/A	N/A	N/A	W0: -0.10 (-0.17; -0.03) P=0.0063, W8: -0.05 (-0.14;0.03) P=0.2293††	CL elicited 2 x as many AEs as O and % of participants with ≥ 1AE significantly greater with CL at W0, however proportions similar at W8 between methods	Between-drug difference in AEs overall (and for frequency of ankle oedema) statistically significant (P=0.02 95% CI 3.1-26.7) for O, not for CL. No difference between methods for other specific AEs or for severity of ankle oedema

Jacobson 1987	Psychiatry	Any AE	NK	106	279	N/A	1871	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	CL elicited 6.7 x the number of AEs than O. Mean number of AEs per assessment 5 more by CL compared to O. In overlapping group of 88 exposed to O, mean number of AEs per assessment was 1.6 x higher than when O was part of combined tool	N/A
Nicholls 1980*	Cardiology	Any AE	8 weeks	24	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	See data by drug arm	Statistically significant difference in mean number of AEs between drugs by CL (5.4 labetolol, 3.6 propranolol P<0.05) but not by O (1.6 labetolol, 1.5 propranol)
O'Connell 2007	Dysmenorrhoea	Any AE	12 weeks	76	66	N/A	177	N/A	N/A	N/A	45(60)	N/A	57(77)	N/A	N/A	N/A	N/A	CL elicited 2.7 x the number of AEs than O	Median number of AEs for contraceptive and placebo by O was 1. By CL median for both was 2
Perez 2012	Parkinson's Disease (PD) and post-stroke controls (PSC)	Any AE	1 occasion	203 PD, 52 PSC	113+6	N/A	PD 1573, PSC 167	N/A	N/A	N/A	PD 85(42), PSC 5(10)	N/A	PD 203(100), PSC 47(90)	N/A	N/A	N/A	N/A	Significantly more participants reporting ≥ 1 AE on CL compared to O (P<0.01) in both groups. Only factor found related to reporting of ≥ 1 AE by participants in response to O, was if participant reported > 2 AEs by CL (OR 1.2(1.1-3.2)	N/A
Rabkin 1992	Psychiatry	Any AE	4 weeks	180/226	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	CL elicited 5 x mean number of AEs than O	N/A

Reilly 1992	Cardiology	Any AE	10 weeks	92	W0 37 W12 46	N/A	W0 340 W12 96	N/A	N/A	N/A	W0 27(29.5) W12 33(35.2)	N/A	WO 84(90.9) W12 74(80.7)	N/A	N/A	N/A	W0: 0-.65 (-0.75;-0.54) P=0.0000, W12: -0.47 (-0.59; -0.34) P=0.0000 ††	% participants with ≥ 1AE was significantly greater with CL at W0 and W12. Only 6% of symptoms reported on CL were reported on O	N/A
Rosenthal 1996	Cardiology	Any AE	12 weeks	5559	984	N/A	7055	N/A	N/A	N/A	705(12.7)	N/A	2753(50)	N/A	N/A	N/A	-0.38 (-0.39;-0.36) P=0.0000††	% of participants with ≥ 1AE was significantly greater with CL	Netween-drug difference in AEs overall was statistically significant (-0.03 (-0.07; -0.001) P=0.0266) for O, not for CL.(-0.04 -0.09; 0.004) P=0.0776
Wallin 1981	Gonorrhoea	Any AE	NK	515	25	N/A	16	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	64% increase in number of AEs elicited by CL after O	N/A
Wernicke 2006	Not known	Any AE	NK	635	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	See data by drug arm	Sp-So index > 1.0 for 22/29 (75.9%) AEs but not significant for most: O more effective in detecting difference between treatments. More statistically significant differences between treatments by CL (9 AEs) than O (5 AEs): differences in % of AEs between drug and placebo (rather than ratios of AE rates) more often greater with CL

de Vries 2014	Diabetes	Any AE	1 occasion	78	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	Sensitivities, PPV of CL compared with O(D) at primary SOC (95% CI): 4-weeks 33% (4-78) and 10% (1-30); 3-months 33% (21-47) and 51% (34-69). Sensitivities at specific ADE level (95% CI): 4-weeks 43% (10-92); 3-months 41% (30-54)	N/A
Sheftell 2004	Migraine	Any AE	1 occasion	415	N/A	N/A	N/A	N/A	N/A	N/A	118 (28.4)	N/A	248 (59.8)	N/A	N/A	N/A	-0.31 (-0.38; -0.25) P=0.0000	Significant more AEs reported through CL for those reporting 1, 2, 3 or more AEs	N/A
Landen 2005	Psychiatry	Sexual	4 weeks	119	N/A	N/A	N/A	N/A	N/A	N/A	7 (6)	N/A	N/A	N/A	49 (41)	N/A	OR 11 (5;26) χ^2 45 p<0.001	R elicited significantly greater number of AEs than O. 2 women versus 5 men reported AE by O (χ^2 6.7 p=0.01) and 29 women versus 20 men by R (χ^2 =3.7 p=0.06)	No statistically significant difference found between drugs by either method
Yeo 1991	Cardiology	Any AE	24 weeks	128	N/A	N/A	N/A	N/A	N/A	N/A	4 (3.1)	N/A	N/A	N/A	12 (20)	N/A	-0.08 (-0.15; -0.01) P=0.0374	% participants with ≥ 1AE significantly greater with R compared to O. Increase in cough by R less consistent in men than women	N/A
Kruft 2007	Ophthalmology	Ocular	NK	NK	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	For 13 of 14 questions there was a statistically greater positive response to CL or R than O	N/A

Montiero 1987	Psychiatry	Sexual	12 weeks	33	N/A	N/A	N/A	N/A	N/A	N/A	3	N/A	23	N/A	N/A	8	N/A	Of 10 participants (all active) who did not report AE by CL, 3 reported AE by O and 8 by INT. 36% of those with drug-induced AE at interview did not report at CL despite concern with it and even if they were secretly reducing dose of drug to overcome it	No AEs in placebo arm by any method
de Vries 2013	Diabetes	Any AE	1 occasion	90	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	Number of AEs similar between the CLs (Z=-0.049, P=0.961)	N/A
Lundberg 1980	Antihistamine	Any AE	9 days	12	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	Two-factor repeated measure of variance for AEs reported by ≥ 50% sample showed significant effects of drug for 6 symptoms by CL but no difference by R
Wallander 1991	Cardiology	Any AE	8 weeks	191/251	N/A	N/A	926	N/A	1521	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	Higher mean number AEs through R versus CL (overall and sex/age groups). In addition to measuring frequency, R also quantified degree of change in symptoms	N/A
Allen 2013*	Malaria	Any AE	3-7 days	18, 80 (sites' data cannot be combined)	6, 23	N/A	+1, +20	N/A	N/A	+0, +1**	N/A	N/A	N/A	N/A	N/A	N/A	N/A	% increase in number of AEs: Site 1: 16.7% O to CL, no change with INT. Site 2: 87.0% O to CL, 2.3% CL to INT (subset)	N/A

Greenhill 2004	Psychiatry	Any AE	1 occasion	59	48	N/A	+16	+129	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	Cumulative % increase in number of AEs: 33% O to first CL (drug-specific inquiry, DSI) followed by 202% to next CL (body system review, BSR)	N/A
Os 1994	Cardiology	Cough	2 weeks	828	N/A	N/A	N/A	N/A	N/A	N/A	48 (6)	N/A	185 (22)	N/A	NK	N/A	-0.17 (-0.20; -0.13) P=0.0000	% participants with ≥ 1AE significantly greater with CL compared to O. Not possible to compare frequency overall with R	Cough more frequent with lisinopril than nifedipine (8.5 vs 3.1%, P=0.0009) by O. Similar change in frequency with R. Not possible to present same data for CL. O: AE 3 x more frequent in female vs male with lisinopril (12.6vs4.4%, P=0.0027), no difference for nifedipine. 3 fold difference with CL. By VAS, participant and spouse assessed frequency of lisinopril-associated AE similarly. With O, similar number of AEs independent of smoking, with CL a statistically significant difference between non-smokers and smokers (16vs7%, P=0.0018).

O: open question/spontaneous, O(B): blank page, CL: checklist, R: rating scale. INT: interview

*Two sites/groups reported separately as different participant populations

**subset of participants from C vs CL comparison then interviewed

† McNemar's test of proportions, †† 2-sample t test

Supplementary Table S3.6c: Effects of methods on the nature of AEs reported for between-participant comparisons (Chapter 3)

Study	Location	Within / outside trial	Participants	Therapy area	Endpoint	Treatment-emergent	Duration of follow-up	Effect (nature of AEs)
Avery 1967	US	Within	Patients	Psychiatry	Any AE	Unclear	5 weeks	A statistically significant higher mean severity AEs at each visit by CL overall and in just the active arm. See paper for details.
Huskisson 1974**	Europe	Within	Patients	Rheumatology	Any AE	Unclear	24 weeks	Severity was included as a composite measure with frequency so presented under effect (number of AEs).
Torok 1984***	Europe	Within	Patients	Cardiology	Any AE	Not necessarily	Various	N/A
Bent 2006	US	Within	Patients	Prostatic hyperplasia	Any AE	Unclear	< 1 - 24 months	N/A
Borghi 1984*	Europe	Within	Patients	Cardiology	Any AE	Unclear	16 weeks	N/A

Spilker 1987	US	Outside	Healthy volunteers	No indication	Any AE	Not necessarily	1 occasion	Most common symptoms by CL were fatigue, headache and nasal congestion and by O were headache, back or muscle pain and nasal congestion.
Ciccolunghi 1975	Europe	Outside	Healthy volunteers	No indication	Any AE	Not necessarily	1 occasion	O was associated with a greater severity of symptoms than CL. The type of symptoms reported did depend to some extent on the method.
Brent 2009**	US	Within	Patients	Psychiatry	Self-harm	Yes	12 weeks	
Suicidal								There was no difference between R and O for reporting of serious suicidal or non-suicidal AEs (8.4% vs 7.3%, χ^2=0.03, df=1, p=0.87). Time to onset for suicidal and non-suicidal AEs earlier for CL than O: median 2 vs 5 weeks (χ^2=9.41, df=1, p=0.004).
Non-suicidal								
Suicide attempts								

O: open question/spontaneous, O(B): blank page, CL: checklist, R: rating scale

* All participants asked AEs by open question (O), but as this process possibly involved 'filtering' of reports by the doctor data not included in the review.

** Huskisson used a composite measure of frequency and severity

*** A selection of the total AEs are presented in review as those objectively measured were excluded

Supplementary Table S3.6d: Effects of methods on the nature of AEs reported for within-participant comparisons (Chapter 3)

Study	Location	Within / outside trial	Participants	Therapy area	Endpoint	Treatment-emergent	Duration of follow-up	Effect (nature of AEs)
Barber 1995	US	Within	Patients	Ophthalmology	Any AE	Unclear	4 weeks	Participants reporting AEs by O and discontinuing therapy reported more bother and activity limitation compared to those who did not report by O but did by CL. Average global QoL scores increased as participants reported AEs by O and discontinued therapy; participants who reported AEs by O indicated more negative impact of side-effects and activity imitations on QoL, more dissatisfaction with medication, and more non-compliance versus those not reporting by O.
Barrowman 1970	Europe	Within	HV	GI	Any AE	Yes	1 occasion	N/A
Downing 1970	US	Within	Patients	Psychiatry	Any AE	Unclear	4 weeks	Greater proportion of participants reporting an AE by both O and CL (26, 90%) had high mean intensity scores compared to those only reporting by CL (12, 50% - predominantly those taking tranquilizers) P<0.01. The former participants more often reported AEs at a high discomfort level compared to the latter P< 0.05.
Hermans 1994	Europe	Within	Patients	Cardiology	Any AE	Yes	6 weeks	There was no apparent difference between the methods for severity and duration of AEs.

Jacobson 1987	US	Within	Patients	Psychiatry	Any AE	Unclear	NK	CL detected a greater variety of AEs than O while AEs reported by O had a higher mean severity compared to those reported by CL.
Nicholls 1980*	Europe	Within	Patients	Cardiology	Any AE	Unclear	8 weeks	N/A
O'Connell 2007	US	Within	Patients	Dysmenorrhoea	Any AE	Unclear	12 weeks	N/A
Perez 2012	Europe	Outside	Patients	Parkinson's Disease, post-stroke controls	Any AE	Unclear	1 occasion	No relationship between AE severity and the O questioning method.
Rabkin 1992	US	Within	Patients	Psychiatry	Any AE	Unclear	4 weeks	AEs reported to O significantly more distressing, more often interfered daily functioning and elicited more changes in clinical management versus CL (NB for latter additional 46 participants assessed). No medically serious AEs elicited by O alone. Overall, and for participants on active drug (but not placebo), mean severity of AEs reported by O significantly greater versus CL. However, 61% of AEs rated severe/very severe elicited by CL,, 65% AEs causing severe/very severe dysfunction detected by CL versus 35% by O.
Reilly 1992	US	Within	Patients	Cardiology	Any AE	Not necessarily	10 weeks	Symptoms reported by O more bothersome than CL. no change in mean degree of distress caused by AEs using CL, but increase in distress associated with AEs by O. Duration of AEs similar for O and CL but a higher symptom severity score by O versus CL. Significant relationship between degree of bother (P<0.0001), duration (P=0.02), severity (0.0003) and reporting of AEs by O. Only 18% of symptoms bothering participants a lot/extremely were first reported by CL.
Rosenthal 1996	Europe	Within	Patients	Cardiology	Any AE	Unclear	12 weeks	N/A
Wallin 1981	Europe	Within	Patients	Gonorrhoea	Any AE	Unclear	NK	Conclusions severity not supported by the data.
Wernicke 2006	Not known	Within	Patients	Not known	Any AE	Yes	NK	N/A

de Vries 2014	Europe	Outside	Patients	Diabetes	Any AE	Unclear	1 occasion	N/A
Sheftell 2004	Multinational	Outside	Patients	Migraine	Any AE	Unclear	1 occasion	No difference between O and CL for severity. However 31 (7.5%) participants who rated AE as severe in CL did not report the AE in O.
Landen 2005	Europe	Within	Patients	Psychiatry	Sexual	Yes	4 weeks	
Yeo 1991	Europe	Within	Patients	Cardiology	Any AE	Unclear	24 weeks	N/A
Kruft 2007	Multinational	Within	Patients	Ophthalmology	Occular	Unclear	NK	N/A
Montiero 1987	Europe	Within	Patients	Psychiatry	Sexual	Yes	12 weeks	N/A
de Vries 2013	Europe	Outside	Patients	Diabetes	Any AE	Unclear	1 occasion	N/A
Lundberg 1980	US	Within	Patients	Antihistamine	Any AE	Unclear	9 days	N/A
Wallander 1991	Europe	Within	Patients	Cardiology	Any AE	Unclear	8 weeks	N/A
Allen 2013*	Africa	Within	Patients, healthy volunteers	Malaria	Any AE	Yes	3-7 days	All additional AEs were mild and unlikely related to trial drug
Greenhill 2004	US	Outside	Patients	Psychiatry	Any AE	Unclear	1 occasion	54% of AEs elicited by O were moderate to severe compared to 75% of those elicited by DSI and 37% for BSR. Of the 17 severe AEs, 6 (37%) were elicited by BSR. 31% of the AEs elicited by O were clinically relevant compared with 12% for the DSI and 15% for the BSR. Of the clinically relevant AEs (n=37), 19 (53%) were elicited by BSR.
Os 1994	Europe	Within	Patients	Cardiology	Cough	Yes	2 weeks	N/A

O: open question/spontaneous, O(B): blank page, CL: checklist, R: rating scale. INT: interview
QOL: quality of life
*Two sites/groups reported separately as different participant populations

Supplementary Table S5.1: Checklist items by body systems, symptoms, diseases and treatments (Chapter 5)

Common to both trials	Additional in South Africa	Additional in Tanzania
Body systems		
Head and neck	Muscles/bones/joints	Endocrine/metabolic
Eyes, ears, nose, throat, mouth	Skin	
Cardiovascular (heart)	Blood	
Pulmonary (lungs)		
Renal (kidney/bladder)		
Gastro intestinal/hepatobiliary (stomach/bowels/liver)		
Symptoms		
Headache	Painful mouth sores	Body pain
Abdominal pain (heartburn/indigestion)	Pain swallowing	Night sweats
Diarrhoea	Constipation	Dizziness
Nausea	Bone aches	Convulsion
Vomiting	Skin infection/rash	Insomnia
Fever	Flu/cold	Problems concentrating Tingling/pain
Weakness	Pain/blood on urination	Palpitations
Fatigue		Chest pain
Muscle, joint aches		
Asthma/breathing problems/dyspnoea		
Skin itch		
Cough		
Anorexia/loss of appetite		
Treatments*		
Antacids, detoxifiers, digestives	Dutch remedies (local South African remedies)	Anti-hypertensives
Laxatives/purgatives/cleansers/enemas		Anti-malarial treatments
Anti-diarrhoeals		
Skin products		
Cough/cold/flu preparations		
Antihistamines/anti-itch treatments		
Antibiotics/anti-infection preparations		
Painkillers/anti-inflammatories		
Worm treatments		
Anti-asthmatics		
Eye, ear or nose drops/treatments		
Vitamins, minerals, iron, supplements		
Immune boosters (including tonics)		
Contraceptives (by mouth or injection)		

*Examples of products from the locality to each trial were also presented to participants as photographs

Supplementary Table S5.2: South African participants' reports of adverse events, medical histories and concomitant medications elicited by question method (Chapter 5)

Medical histories

System organ class†	Preferred term†	Number of reports by general enquiry	Additional number of reports by checklists	Additional number of reports by in-depth interview
Gastrointestinal disorders	Constipation, dental caries, diarrhoea	1	1	1
General disorders	Malaise, fatigue	0	4	0
Injury, poisoning and procedural complications	Gunshot wound	1	0	0
Metabolism and nutrition disorders	Decreased appetite	0	1	0
Musculoskeletal and connective tissue disorders	Myalgia	0	1	0
Nervous system disorders	Amnesia, insomnia	0	0	2
Reproductive system and breast disorders	Dysmenorrhoea	1	0	0
Respiratory, thoracic and mediastinal disorders	Sinusitis (chronic)	0	1	0
Skin and subcutaneous tissue disorders	Dermatitis	0	0	1
Vascular disorders	Flushing	1	0	0
Totals		**4**	**8**	**4**

†Medical Dictionary for Regulatory Authorities (MedDRA®) terminology

Adverse events

System organ class†	Preferred term†	Number of reports by general enquiry	Additional number of reports by checklists	Additional number of reports by in-depth interview
Eye disorders	Conjunctivitis, uveitis, eye pruritus	2	1	0
Gastrointestinal disorders	Diarrhoea, vomiting, abdominal pain, constipation, nausea	4	3	0
General disorders	Catheter site haematoma, pyrexia, chest discomfort, mucosal inflammation, lethargy, night sweats	2	3	0
Immune system disorders	Seasonal allergy	1	0	0
Infections and infestations	Herpes zoster, molluscum contagious, Acarodermatitis	3	0	0
Injury, poisoning and procedural complications	Arthropod bites	1	1	0
Investigations	Weight decreased	1	0	0
Metabolism and nutrition disorders	Decreased appetite	1	1	0
Musculoskeletal and connective tissue disorders	Neck pain, back pain, arthralgia	0	4	0
Nervous system disorders	Headache	2	2	0
Psychiatric disorders	Self-induced vomiting	0	1	0
Renal and urinary disorder	Polyuria	1	1	0
Reproductive system and breast disorders	Breast pain	1	0	0
Respiratory, thoracic and mediastinal disorders	Influenza, nasal discomfort and throat irritation	4	2	0
Skin and subcutaneous tissue disorders	Acne	0	1	1
Totals		**23**	**20**	**1**

†Medical Dictionary for Regulatory Authorities (MedDRA®) terminology

Medications

Anatomical Therapeutic Chemical (ATC) classification	Product names	Number of reports by general enquiry	Additional number of reports by checklists	Additional number of reports by in-depth interview
A02 DRUGS FOR ACID RELATED DISORDERS	Gelusil (Aluminium hydroxide, Magnesium trisilicate)	0	1	0
A03 DRUGS USED FOR FUNCTIONAL GASTROINTESTINAL DISORDERS	Millerspas (hyoscine), Vinegar (as emetic)	1	1	0
A06 LAXATIVES	Sorbitol (sorbose), Magnesium sulphate	1	2	0
A07 ANTIDIARRHEALS, INTESTINAL ANTIINFLAMMATORY/ANTIINFECTIVE AGENTS	Anti-diarrhoeal (UNK)	0	1	0
A11 VITAMINS	Vitamins	1	1	0
D02 EMOLLIENTS AND PROTECTIVES	Aqueous cream	0	2	1
D07 CORTICOSTEROIDS, DERMATOLOGICAL PREPARATIONS	Hydrocortisone cream	0	0	2
G03 SEX HORMONES AND MODULATORS OF THE GENITAL SYSTEM	Petogen (contraceptive)	0	1	0
J01 ANTIBACTERIALS FOR SYSTEMIC USE	Doxycycline, Erythromycin	2	0	0
J05 ANTIVIRALS FOR SYSTEMIC USE	Acyclovir	1	0	0
N02 ANALGESICS	Paracetamol, GrandPa (aspirin, paracetamol, caffeine), Myprodol (ibuprofen, paracetamol, codeine)	2	6	0
N06 PSYCHOANELEPTICS	Amitriptyline	1	0	0
P02 ANTHELMINTICS	D-worm (mebendazole), Anthelminthic (UNK)	1	3	0
P03 ECTOPARASITICIDES, INCL. SCABICIDES, INSECTICIDES AND REPELLENTS	Insect repellent cream	1	0	0
R03 DRUGS FOR OBSTRUCTIVE AIRWAY DISEASES	Alcophyllex (theophylline, diphenhydramine, ammonium chloride)	1	0	0
R05 COUGH AND COLD PREPARATIONS	Flu remedy (paracetamol, antihistamines, caffeine decongestants, vitamin C)	3	0	1

R06 ANTIHISTAMINES FOR SYSTEMIC USE	Chlorpheniramine	1	0	0
S01 OPTHALMOLOGICALS	Spersallerg eye drops (tetryzoline, antazoline)	1	1	0
Unclassified (herbal)	Herbal ointment (UNK), Herbal tea (sutherlandia), Herbal syrup (aloe)	0	3	0
Unclassified (used orally)	Camphor	0	1	0
Total		**17**	**23**	**4**

Supplementary Table S5.3: Tanzanian participants' reports of adverse events, medical histories and concomitant medications elicited by question method (Chapter 5)

Medical histories

System organ class†	Preferred term†	Number of reports by general enquiry	Additional number of reports by checklists	Additional number of reports by in-depth interview
Cardiac disorders	Palpitations	5	9	2
Eye disorders	Visual acuity reduced	2	0	1
Gastrointestinal disorders	Abdominal pain, constipation, diarrhoea, mouth ulceration, nausea, abdominal discomfort, abdominal distension, peptic ulcer, toothache, vomiting, faeces discoloured	25	20	4
General disorders	Fatigue, pyrexia, pain, granuloma, stretching, night sweats, asthenia, malaise	68	126	3
Hepatobiliary disorders	Cholecystitis	1	0	0
Infections and infestations	Appendicitis, furuncle, herpes zoster, onychpmycosis, ear infection, periodontitis, schistosomiasis, tuberculosis	22	6	1
Investigations	Weight decreased	0	4	0
Metabolism and nutrition disorders	Decreased appetite,	6	6	0
Musculoskeletal and connective tissue disorders	Back pain, neck pain, pain in extremity, myalgia, arthralgia	31	14	1
Neoplasms benign, malignant and unspecified (incl. cysts and polyps)	Uterine leiomyoma	1	0	0
Nervous system disorders	Dizziness, headache, hypoaesthesia, parathesia, burning sensation	62	32	0
Pregnancy, puerperium and perinatal conditions	Abortion	1	0	0
Renal and urinary disorder	Dysuria, nephrolithiasis	3	0	0
Reproductive system and breast disorders	Breast enlargement, infertility, pelvic pain, vaginal prolapse, vaginal discharge	9	0	0

Respiratory, thoracic and mediastinal disorders	Asthma, cough, dyspnoea, chest pain, rhinorrhoea	40	21	3
Skin and subcutaneous tissue disorders	Skin hypopigmentation, rash pruritic, skin lesion	8	7	0
Vascular disorders	Hypertension	1	0	0
Totals		**285**	**245**	**15**

†Medical Dictionary for Regulatory Authorities (MedDRA®) terminology

Adverse events

System organ class†	Preferred term†	Number of reports by general enquiry	Additional number of reports by checklists	Additional number of reports by in-depth interview
General disorders and administration site conditions	Pyrexia, fatigue	1	1	0
Metabolism and nutrition disorders	Decreased appetite	1	0	0
Respiratory, thoracic and mediastinal disorders	Cough, rhinitis	2	0	0
Skin and subcutaneous tissue disorders	Rash	2	0	0
Totals		**6**	**1**	**0**

†Medical Dictionary for Regulatory Authorities (MedDRA®) terminology

Medications

Anatomical Therapeutic Chemical (ATC) classification	Product names	Number of reports by general enquiry	Additional number of reports by checklists	Additional number of reports by in-depth interview
A02 DRUGS FOR ACID RELATED DISORDERS	Omeprezole	3	0	0
A06 LAXATIVES	Bisacodyl	1	0	0
A11 VITAMINS	Pyridoxine, pyridoxine and thiamine	8	0	0
B03 ANTIANAEMIC PREPARATIONS	Iron plus vitamins and minerals	2	0	0
C03 DIURETICS	Bendroflumethazide	3	0	0
C08 CALCIUM CHANNEL BLOCKERS	Nifedipine	2	0	0
D01 ANTIFUNGALS FOR DERMATOLOGICAL USE	Griseofulvin, salicylic acid/antifungal combinations	3	0	0
D04 ANTIPRURITICS, INCLU. ANTIHISTAMINES, ANESTHETICS, ETC.	Calamine	3	0	0
D07 CORTICOSTEROIDS, DERMATOLOGICAL PREPARATIONS	Hydrocortisone and gentamicin	2	0	0
G01 GYNECOLOGICAL ANTIINFECTIVES AND ANTISEPTICS	Butoconazole	1	0	0
J01 ANTIBACTERIALS FOR SYSTEMIC USE	Amoxycillin, ciprofloxacin, cloxacillin, sulfamethoxazole and trimethoprim, doxycycline, erthythromycin, metronidazole, phenoxypenicillin	66	1	0
J04 ANTIMYCOBACTERIALS	Rifampicin, pyrazinamide, ethambutol, isoniazid	3	0	0
J05 ANTIVIRALS FOR SYSTEMIC USE	Acyclovir	1	0	0
M01 ANTIINFLAMMATORY AND ANTIRHEUMATIC PRODUCTS	Ibuprofen, diclofenac, diclofenac and paracetamol	8	0	1
N02 ANALGESICS	Acetylsalicylic acid, caffeine and paracetamol, aspirin and caffeine, paracetamol, “analgesic”	63	1	5
P03 ECTOPARASITICIDES, INCL. SCABICIDES, INSECTICIDES AND REPELLENTS	Benzyl benzoate	1	0	0
N06 PSYCHOANELEPTICS	Amitriptyline	3	0	0

P01 ANTIPROTOZOALS	Artemether and lumefantrine, sulfadoxine and pyrimethamine	4	0	1
P02 ANTHELMINTICS	Albendazole	1	0	0
R03 DRUGS FOR OBSTRUCTIVE AIRWAY DISEASES	Aminophylline, salbutamol	4	0	0
R06 ANTIHISTAMINES FOR SYSTEMIC USE	Cetirizine, chlorphenamine, promethazine	8	0	0
Unclassified (herbal)		6	0	2
Total		**196**	**2**	**9**

Supplementary File S3.1: Systematic review search strategies (Chapter 3)

Database(s): Embase 1980 to 2013 Week 17, Embase 1980 to 2013 Week 27 Searched April 2013

16th July 2013 (used entry week field and selected all from 201317 to latest 201328) – 158 results when limited to human and English

16th March 2015 used Entry field 201327 to 2015 wk 11 retrieved 1405

#	Searches	Results
1	exp adverse drug reaction/	303737
2	drug safety/	210956
3	side effect/	157758
4	(drug? adj2 (safety or harm? or adverse or undesirable or tolerability or toxicity or toxic or effect? or event? or symptom?)).ti,ab.	66980
5	((adverse or side or undesirable or treatment emergent or treatment related) adj2 (effect? or reaction? or event? or outcome? or symptom?)).ti,ab.	504926
6	(adr or adrs).ti,ab.	8291
7	(drug? adj2 (safety or harm? or adverse or undesirable or tolerability or toxicity or toxic or effect? or event? or symptom?) adj3 (measur$ or assess$ or monitor$ or detect$ or report$ or self report$ or record$ or identif$ or collect$ or notify or notifie? or notification$)).ti,ab.	7902
8	((adverse or side or undesirable or treatment emergent or treatment related) adj2 (effect? or reaction? or event? or outcome? or symptom?) adj3 (measur$ or assess$ or monitor$ or detect$ or report$ or self report$ or record$ or identif$ or collect$ or notify or notifie? or notification$)).ti,ab.	53530
9	((adr or adrs) adj3 (measur$ or assess$ or monitor$ or detect$ or report$ or self report$ or record$ or identif$ or collect$ or notify or notifie? or notification$)).ti,ab.	2377
10	or/1-9	935536
11	adverse.ab. /freq=3 or adr.ab. /freq=3 or adrs.ab. /freq=3 or ae.ab. /freq=3 or aes.ab. /freq=3 or (side adj effect$).ab. /freq=3 or elicit$.ab. /freq=3 or symptom?.ab. /freq=3	192692
12	10 and 11	62830
13	(drug? adj2 (safety or harm? or adverse or undesirable or tolerability or toxicity or toxic or effect? or event? or symptom?) adj3 (measur$ or assess$ or monitor$ or detect$ or report$ or self report$ or record$ or identif$ or collect$ or notify or notifie? or	1952

	notification$)).ti.	
14	((adverse or side or undesirable or treatment emergent or treatment related) adj2 (effect? or reaction? or event? or outcome? or symptom?) adj3 (measur$ or assess$ or monitor$ or detect$ or report$ or self report$ or record$ or identif$ or collect$ or notify or notifie? or notification$)).ti.	3406
15	((adr or adrs) adj3 (measur$ or assess$ or monitor$ or detect$ or report$ or self report$ or record$ or identif$ or collect$ or notify or notifie? or notification$)).ti.	159
16	or/12-15	65507
17	((patient$ or participant$ or subject$) adj2 (elicit$ or evoke$ or solicit$ or unsolicit$ or notify or notifie? or notification$ or spontaneous$ or prompt or prompted or unprompted or open-ended or structured or systematic or standardi$)).ti,ab.	22183
18	(spontaneous report$ or self report$ or participant report$ or patient report$ or subject report$ or self administer$).ti,ab.	129006
19	((patient$ or participant$ or subject$) adj2 (enquir$ or inquir$ or complain$ or checklist$ or check-list$ or query or querie$ or form or forms or interview$)).ti,ab.	49399
20	((elicit$ or evoke$ or solicit$ or unsolicit$ or notify or notification$ or prompt or prompted or unprompted or open-ended or structured or systematic or standardi$) adj2 (enquir$ or inquir$ or complain$ or checklist$ or check-list$ or query or querie$ or form or forms or interview$)).ti,ab.	41109
21	or/17-20	231691
22	and/16,21	4428

(limit to English and Human gives 3478)

Database(s): Ovid MEDLINE(R) 1946 to April Week 3 2013, MEDLINE(R) 1946 to July Week 1 2013
Searched April 2013

16th July 2013 (searched 201304$.ed,ep,up) – 63 results when limited to Humans and English

16th March 2015 searched to 2015 March wk 2 retrieved 702

#	Searches	Results
1	Adverse Drug Reaction Reporting Systems/	5208
2	Drug Toxicity/	5938
3	(ae or de).fs.	3394970

4	(drug? adj2 (safety or harm? or adverse or undesirable or tolerability or toxicity or toxic or effect? or event? or symptom?)).ti,ab.	47522
5	((adverse or side or undesirable or treatment emergent or treatment related) adj2 (effect? or reaction? or event? or outcome? or symptom?)).ti,ab.	350647
6	(adr or adrs).ti,ab.	5344
7	(drug? adj2 (safety or harm? or adverse or undesirable or tolerability or toxicity or toxic or effect? or event? or symptom?) adj3 (measur$ or assess$ or monitor$ or detect$ or report$ or self report$ or record$ or identif$ or collect$ or notify or notifie? or notification$)).ti,ab.	5193
8	((adverse or side or undesirable or treatment emergent or treatment related) adj2 (effect? or reaction? or event? or outcome? or symptom?) adj3 (measur$ or assess$ or monitor$ or detect$ or report$ or self report$ or record$ or identif$ or collect$ or notify or notifie? or notification$)).ti,ab.	35078
9	((adr or adrs) adj3 (measur$ or assess$ or monitor$ or detect$ or report$ or self report$ or record$ or identif$ or collect$ or notify or notifie? or notification$)).ti,ab.	1071
10	or/1-9	3567892
11	adverse.ab. /freq=3 or adr.ab. /freq=3 or adrs.ab. /freq=3 or ae.ab. /freq=3 or aes.ab. /freq=3 or (side adj effect$).ab. /freq=3 or elicit$.ab. /freq=3 or symptom?.ab. /freq=3	132190
12	and/10-11	54020
13	(drug? adj2 (safety or harm? or adverse or undesirable or tolerability or toxicity or toxic or effect? or event? or symptom?) adj3 (measur$ or assess$ or monitor$ or detect$ or report$ or self report$ or record$ or identif$ or collect$ or notify or notifie? or notification$)).ti.	1253
14	((adverse or side or undesirable or treatment emergent or treatment related) adj2 (effect? or reaction? or event? or outcome? or symptom?) adj3 (measur$ or assess$ or monitor$ or detect$ or report$ or self report$ or record$ or identif$ or collect$ or notify or notifie? or notification$)).ti.	2241
15	((adr or adrs) adj3 (measur$ or assess$ or monitor$ or detect$ or report$ or self report$ or record$ or identif$ or collect$ or notify or notifie? or notification$)).ti.	37
16	or/12-15	55859
17	((patient$ or participant$ or subject$) adj2 (elicit$ or evoke$ or solicit$ or unsolicit$ or notify or notifie? or notification$ or spontaneous$ or prompt or prompted or unprompted or open-ended or structured or systematic or standardi$)).ti,ab.	16381

18	(spontaneous report$ or self report$ or participant report$ or patient report$ or subject report$ or self administer$).ti,ab.	97507
19	((patient$ or participant$ or subject$) adj2 (enquir$ or inquir$ or complain$ or checklist$ or check-list$ or query or querie$ or form or forms or interview$)).ti,ab.	34835
20	((elicit$ or evoke$ or solicit$ or unsolicit$ or notify or notification$ or prompt or prompted or unprompted or open-ended or structured or systematic or standardi$) adj2 (enquir$ or inquir$ or complain$ or checklist$ or check-list$ or query or querie$ or form or forms or interview$)).ti,ab.	30608
21	or/17-20	172513
22	16 and 21	3699

(limit to English and Human gives 3397)

Database(s): Ovid MEDLINE(R) In-Process & Other Non-Indexed Citations July 15, 2013 Searched 16th July 2013

#	Searches	Results
1	(drug? adj2 (safety or harm? or adverse or undesirable or tolerability or toxicity or toxic or effect? or event? or symptom?)).ti,ab.	2931
2	((adverse or side or undesirable or treatment emergent or treatment related) adj2 (effect? or reaction? or event? or outcome? or symptom?)).ti,ab.	24324
3	(adr or adrs).ti,ab.	401
4	(drug? adj2 (safety or harm? or adverse or undesirable or tolerability or toxicity or toxic or effect? or event? or symptom?) adj3 (measur$ or assess$ or monitor$ or detect$ or report$ or self report$ or record$ or identif$ or collect$ or notify or notifie? or notification$)).ti,ab.	402
5	((adverse or side or undesirable or treatment emergent or treatment related) adj2 (effect? or reaction? or event? or outcome? or symptom?) adj3 (measur$ or assess$ or monitor$ or detect$ or report$ or self report$ or record$ or identif$ or collect$ or notify or notifie? or notification$)).ti,ab.	2684
6	((adr or adrs) adj3 (measur$ or assess$ or monitor$ or detect$ or report$ or self report$ or record$ or identif$ or collect$ or notify or notifie? or notification$)).ti,ab.	155
7	or/1-6	25915
8	adverse.ab. /freq=3 or adr.ab. /freq=3 or adrs.ab. /freq=3 or ae.ab. /freq=3 or aes.ab. /freq=3 or (side adj effect$).ab. /freq=3 or elicit$.ab. /freq=3 or symptom?.ab. /freq=3	9705

9	and/7-8	2853
10	(drug? adj2 (safety or harm? or adverse or undesirable or tolerability or toxicity or toxic or effect? or event? or symptom?) adj3 (measur$ or assess$ or monitor$ or detect$ or report$ or self report$ or record$ or identif$ or collect$ or notify or notifie? or notification$)).ti.	112
11	((adverse or side or undesirable or treatment emergent or treatment related) adj2 (effect? or reaction? or event? or outcome? or symptom?) adj3 (measur$ or assess$ or monitor$ or detect$ or report$ or self report$ or record$ or identif$ or collect$ or notify or notifie? or notification$)).ti.	196
12	((adr or adrs) adj3 (measur$ or assess$ or monitor$ or detect$ or report$ or self report$ or record$ or identif$ or collect$ or notify or notifie? or notification$)).ti.	7
13	or/9-12	2975
14	((patient$ or participant$ or subject$) adj2 (elicit$ or evoke$ or solicit$ or unsolicit$ or notify or notifie? or notification$ or spontaneous$ or prompt or prompted or unprompted or open-ended or structured or systematic or standardi$)).ti,ab.	990
15	(spontaneous report$ or self report$ or participant report$ or patient report$ or subject report$ or self administer$).ti,ab.	7777
16	((patient$ or participant$ or subject$) adj2 (enquir$ or inquir$ or complain$ or checklist$ or check-list$ or query or querie$ or form or forms or interview$)).ti,ab.	2244
17	((elicit$ or evoke$ or solicit$ or unsolicit$ or notify or notification$ or prompt or prompted or unprompted or open-ended or structured or systematic or standardi$) adj2 (enquir$ or inquir$ or complain$ or checklist$ or check-list$ or query or querie$ or form or forms or interview$)).ti,ab.	2547
18	or/14-17	13045
19	and/13,18	219

CINAHL 1980 to April 2013

16th July 2013 searched 201304 in EM field gave 3 extra results

16th March 2015 searched 201304 to 2015* retrieved 154

Search ID#	Search Terms	Actions

S21	S11 AND S20	True S21	870
S20	S12 OR S13 OR S14 OR S15 OR S16 OR S17 OR S18 OR S19	True S20	58,723
S19	AB ((elicit* or evoke* or solicit* or unsolicit* or notify or notification* or prompt or prompted or unprompted or open-ended or structured or systematic or standardi*) N2 (enquir* or inquir* or complain* or checklist* or check-list* or query or querie* or form or forms or interview*))	True S19	15,094
S18	TI ((elicit* or evoke* or solicit* or unsolicit* or notify or notification* or prompt or prompted or unprompted or open-ended or structured or systematic or standardi*) N2 (enquir* or inquir* or complain* or checklist* or check-list* or query or querie* or form or forms or interview*))	True S18	200
S17	AB ((patient* or participant* or subject*) N2 (enquir* or inquir* or complain* or checklist* or check-list* or query or querie* or form or forms or interview*))	True S17	12,113
S16	TI ((patient* or participant* or subject*) N2 (enquir* or inquir* or complain* or checklist* or check-list* or query or querie* or form or forms or interview*))	True S16	747
S15	AB ("spontaneous report*" or "self report*" or "participant report*" or "patient report*" or "subject report*" or "self administer*")	True S15	28,140
S14	TI ("spontaneous report*" or "self report*" or "participant report*" or "patient report*" or "subject report*" or "self administer*")	True S14	4,309
S13	AB ((patient* or participant* or subject*) N2 (elicit* or evoke* or solicit* or unsolicit* or notify or notifie# or notification* or	True S13	4,297

	spontaneous* or prompt or prompted or unprompted or open-ended or structured or standardi*))	
S12	TI ((patient* or participant* or subject*) N2 (elicit* or evoke* or solicit* or unsolicit* or notify or notifie# or notification* or spontaneous* or prompt or prompted or unprompted or open-ended or structured or standardi*))	True S12 548
S11	S1 OR S2 OR S3 OR S4 OR S5 OR S6 OR S7 OR S8 OR S9 OR S10	True S11 16,075
S10	AB ((adr or adrs) N3 (measur* or assess* or monitor* or detect* or report* or self report* or record* or identif* or collect* or notify or notifie# or notification*))	True S10 125
S9	TI ((adr or adrs) N3 (measur* or assess* or monitor* or detect* or report* or self report* or record* or identif* or collect* or notify or notifie# or notification*))	True S9 7
S8	AB ((adverse or side or undesirable or treatment emergent or treatment related) N2 (effect# or reaction# or event# or outcome# or symptom#) N3 (measur* or assess* or monitor* or detect* or report* or self report* or record* or identif* or collect* or notify or notifie# or notification*))	True S8 6,586
S7	TI ((adverse or side or undesirable or treatment emergent or treatment related) N2 (effect# or reaction# or event# or outcome# or symptom#) N3 (measur* or assess* or monitor* or detect* or report* or self report* or record* or identif* or collect* or notify or notifie# or notification*))	True S7 799
S6	AB (drug# N2 (safety or harm# or adverse or undesirable or tolerability or toxicity or toxic or effect# or event# or symptom#) N3 (measur* or assess* or monitor* or detect* or report* or	True S6 718

	self report* or record* or identif* or collect* or notify or notifie# or notification*))	
S5	TI (drug# N2 (safety or harm# or adverse or undesirable or tolerability or toxicity or toxic or effect# or event# or symptom#) N3 (measur* or assess* or monitor* or detect* or report* or self report* or record* or identif* or collect* or notify or notifie# or notification*))	True S5 364
S4	TI (adr or adrs)	True S4 40
S3	TI ((adverse or side or undesirable or treatment emergent or treatment related) N2 (effect# or reaction# or event# or outcome# or symptom#))	True S3 6,607
S2	TI (drug# N2 (safety or harm# or adverse or undesirable or tolerability or toxicity or toxic or effect# or event# or symptom#))	True S2 2,176
S1	(MH "Adverse Drug Event")	True S1 3,161

Web of Knowledge strategies

Databases=SCI-EXPANDED, SSCI, CPCI-S, BIOSIS Timespan=All years

(numbers are from SCI/SSCI/CPCI-S)

Searched 26/6/13

Repeated 16th July 2013, re run search with “Records process from” 2013-05-01, retrieved 3 results from Web of Science databases and 1 from BIOSIS (latter can only do 2013 so may be duplicate)

16th March 2015 pubn date 2013-5 retrieved 271

#13 241 #11 OR #9

12 170 #11 NOT #9

11 241 #10 AND #8 AND #4

10 11,417,480 TS=(measur* or assess* or monitor* or detect* or report* or self report* or record* or identif* or collect* or notify or notified or notification*)

9 16 #8 AND #5

8 52,126 #7 OR #6

7 46,392 TS=(patient* NEAR/2 enquir*) or TS=(participant* NEAR/2 enquir*) or TS=(subject* NEAR/2 enquir*) or TS=(patient* NEAR/2 inquir*) or TS=(participant* NEAR/2 inquir*) or TS=(subject* NEAR/2 inquir*) or TS=(patient* NEAR/2 complain*) or TS=(participant* NEAR/2 complain*) or TS=(subject* NEAR/2 complain*) or TS=(patient* NEAR/2 checklist*) or TS=(participant* NEAR/2 checklist*) or TS=(subject* NEAR/2 checklist*) or TS=(patient* NEAR/2 check-list*) or TS=(participant* NEAR/2 check-list*) or TS=(subject* NEAR/2 check-list*) or TS=(patient* NEAR/2 query) or TS=(participant* NEAR/2 query) or TS=(subject* NEAR/2 query) or TS=(patient* NEAR/2 querie*) or TS=(participant* NEAR/2 querie*) or TS=(subject* NEAR/2 querie*) or TS=(patient* NEAR/2 form) or TS=(participant* NEAR/2 form) or TS=(subject* NEAR/2 form) or TS=(patient* NEAR/2 forms) or TS=(participant* NEAR/2 forms) or TS=(subject* NEAR/2 forms) or TS=(patient* NEAR/2 interview*) or TS=(participant* NEAR/2 interview*) or TS=(subject* NEAR/2 interview*)

6 5,972 TS=(patient* NEAR/2 elicit*) or TS=(participant* NEAR/2 elicit*) or TS=(subject* NEAR/2 elicit*) or TS=(patient* NEAR/2 evoke*) or TS=(participant* NEAR/2 evoke*) or TS=(subject* NEAR/2 evoke*) or TS=(patient* NEAR/2 solicit*) or TS=(participant* NEAR/2 solicit*) or TS=(subject* NEAR/2 solicit*) or TS=(patient* NEAR/2 unsolicit*) or TS=(participant* NEAR/2 unsolicit*) or TS=(subject* NEAR/2 unsolicit*) or TS=(patient* NEAR/2 notif*) or TS=(participant* NEAR/2 notif*) or TS=(subject* NEAR/2 notif*) or TS=(patient* NEAR/2 prompted) or TS=(participant* NEAR/2 prompted) or TS=(subject* NEAR/2 prompted) or TS=(patient* NEAR/2 unprompted) or TS=(participant* NEAR/2 unprompted) or TS=(subject* NEAR/2 unprompted)

5 5,551 #4 AND #3

4 48,150 #2 OR #1

3 2,594,464 TI=(measur* or assess* or monitor* or detect* or report* or self report* or record* or identif* or collect* or notify or notified or notification*)

2 30,484 TI=(“adverse effect*” or “side effect*” or “undesirable effect*” or “adverse reaction*” or “side reaction*” or “undesirable reaction*” or “adverse event*” or "undesirable event*” or “adverse outcome*” or “undesirable outcome*” or “adverse symptom*” or “undesirable symptom*” or “treatment emergent” or “treatment related”)

1 19,049 TI=(drug* NEAR/2 safety) or TI= (drug* NEAR/2 harm*) or TI= (drug* NEAR/2 adverse) or TI= (drug* NEAR/2 undesirable) or TI= (drug* NEAR/2 tolerability) or TI= (drug* NEAR/2 toxicity) or TI= (drug* NEAR/2 toxic) or TI= (drug* NEAR/2 effect) or TI= (drug* NEAR/2 effecTI) or TI= (drug* NEAR/2 event) or TI= (drug* NEAR/2 evenTI) or TI= (drug* NEAR/2 symptom) or TI= (drug* NEAR/2 symptoms)

Database(s): CAB Abstracts 1973 to 2013 Week 24
Searched 25/6/13

Searched 16th Jul 2013 (used update code field and selected all for June and July), but no new records

Searched 16th March 2015 update code to 2015 wk10 retrieved 45

#	Searches	Results
1	adverse effects/	27215
2	drug toxicity/	6475
3	(drug? adj2 (safety or harm? or adverse or undesirable or tolerability or toxicity or toxic or effect? or event? or symptom?)).ti,ab.	4157
4	((adverse or side or undesirable or treatment emergent or treatment related) adj2 (effect? or reaction? or event? or outcome? or symptom?)).ti,ab.	67499
5	(adr or adrs).ti,ab.	633
6	(drug? adj2 (safety or harm? or adverse or undesirable or tolerability or toxicity or toxic or effect? or event? or symptom?) adj3 (measur$ or assess$ or monitor$ or detect$ or report$ or self report$ or record$ or identif$ or collect$ or notify or notifie? or notification$)).ti,ab.	404
7	((adverse or side or undesirable or treatment emergent or treatment related) adj2 (effect? or reaction? or event? or outcome? or symptom?) adj3 (measur$ or assess$ or monitor$ or detect$ or report$ or self report$ or record$ or identif$ or collect$ or notify or notifie? or notification$)).ti,ab.	4292
8	((adr or adrs) adj3 (measur$ or assess$ or monitor$ or detect$ or report$ or self report$ or record$ or identif$ or collect$ or notify or notifie? or notification$)).ti,ab.	87
9	or/1-8	88969
10	adverse.ab. /freq=3 or adr.ab. /freq=3 or adrs.ab. /freq=3 or ae.ab. /freq=3 or aes.ab. /freq=3 or (side adj effect$).ab. /freq=3 or elicit$.ab. /freq=3 or symptom?.ab. /freq=3	20765
11	9 and 10	3200
12	(drug? adj2 (safety or harm? or adverse or undesirable or tolerability or toxicity or toxic or effect? or event? or symptom?) adj3 (measur$ or assess$ or monitor$ or detect$ or report$ or self report$ or record$ or identif$ or collect$ or notify or notifie? or notification$)).ti.	83
13	((adverse or side or undesirable or treatment emergent or treatment related) adj2 (effect? or reaction? or event? or outcome? or symptom?) adj3 (measur$ or assess$ or monitor$ or detect$ or report$ or self report$ or record$ or identif$ or collect$ or notify or notifie? or notification$)).ti.	203

14	((adr or adrs) adj3 (measur$ or assess$ or monitor$ or detect$ or report$ or self report$ or record$ or identif$ or collect$ or notify or notifie? or notification$)).ti.	1
15	or/11-14	3359
16	((patient$ or participant$ or subject$) adj2 (elicit$ or evoke$ or solicit$ or unsolicit$ or notify or notifie? or notification$ or spontaneous$ or prompt or prompted or unprompted or open-ended or structured or systematic or standardi$)).ti,ab.	925
17	(spontaneous report$ or self report$ or participant report$ or patient report$ or subject report$ or self administer$).ti,ab.	12019
18	((patient$ or participant$ or subject$) adj2 (enquir$ or inquir$ or complain$ or checklist$ or check-list$ or query or querie$ or form or forms or interview$)).ti,ab.	2987
19	((elicit$ or evoke$ or solicit$ or unsolicit$ or notify or notification$ or prompt or prompted or unprompted or open-ended or structured or systematic or standardi$) adj2 (enquir$ or inquir$ or complain$ or checklist$ or check-list$ or query or querie$ or form or forms or interview$)).ti,ab.	6044
20	or/16-19	21388
21	15 and 20	167
22	from 21 keep 2, 4, 7-9, 11-12, 15, 17...	75
23	limit 22 to english language	71

Cochrane Library – ran the same strategy and downloaded results from CMR and HTA, then altered to :ti only for AE terms and downloaded CCTR (since I don't have the option of reducing the numbers with freq operator in Cochrane).

Limited to 2013 for update search (16th Jul 2013); no results from CMR,

16th March 2015 publication date 2013-5

CMR

Search Name: Elicitation of AE CMR sensitive

Last Saved: 02/05/2013 09:40:00.932

Description:

16th March 2015 re-ran pubn date 2013-5 no results from CMR (don't think it's being updated actually in which case don't need to report this)

ID Search

#1 MeSH descriptor: [Adverse Drug Reaction Reporting Systems] this term only

#2 MeSH descriptor: [Drug Toxicity] this term only

#3 Any MeSH descriptor with qualifier(s): [Adverse effects - AE, Drug effects - DE]

#4 ((drug? near/2 safety) or (drug? near/2 harm?) or (drug? near/2 adverse) or (drug? near/2 undesirable) or (drug? near/2 tolerability) or (drug? near/2 toxicity) or (drug? near/2 effect?) or (drug? near/2 event?) or (drug? near/2 symptom?)):ti,ab

#5 ((adverse near/2 effect?) or (adverse near/2 reaction?) or (adverse near/2 event?) or (adverse near/2 outcome?) or (adverse near/2 symptom?)):ti,ab

#6 ((side near/2 effect?) or (side near/2 reaction?) or (side near/2 event?) or (side near/2 outcome?) or (side near/2 symptom?)):ti,ab

#7 ((undesirable near/2 effect?) or (undesirable near/2 reaction?) or (undesirable near/2 event?) or (undesirable near/2 outcome?) or (undesirable near/2 symptom?)):ti,ab

#8 (("treatment emergent" near/2 effect?) or ("treatment emergent" near/2 reaction?) or ("treatment emergent" near/2 event?) or ("treatment emergent" near/2 outcome?) or ("treatment emergent" near/2 symptom?)):ti,ab

#9 (("treatment related" near/2 effect?) or ("treatment related" near/2 reaction?) or ("treatment related" near/2 event?) or ("treatment related" near/2 outcome?) or ("treatment related" near/2 symptom?)):ti,ab

#10 ("adverse effect*" near/3 assess*) or ("adverse effect*" near/3 measur*) or ("adverse effect*" near/3 detect*) or ("adverse effect*" near/3 notify) or ("adverse effect*" near/3 notification*):ti,ab

#11 ("adverse event*" near/3 assess*) or ("adverse event *" near/3 measur*) or ("adverse event *" near/3 detect*) or ("adverse event *" near/3 notify) or ("adverse event *" near/3 notification*):ti,ab

#12 ("adverse reaction*" near/3 assess*) or ("adverse reaction *" near/3 measur*) or ("adverse reaction *" near/3 detect*) or ("adverse reaction *" near/3 notify) or ("adverse reaction *" near/3 notification*):ti,ab

#13 ("side effect*" near/3 assess*) or ("side effect*" near/3 measur*) or ("side effect*" near/3 detect*) or ("side effect*" near/3 notify) or ("side effect*" near/3 notification*):ti,ab

#14 #1 or #2 or #3 or #4 or #5 or #6 or #7 or #8 or #9 or #10 or #11 or #12 or #13

#15 ((elicit*NEAR/2 questionnaire*) or (evoke* near/2 questionnaire*) or (solicit* near/2 questionnaire*) or (unsolicit* near/2 questionnaire*) or (self-report* near/2 questionnaire*) or (participant-report* near/2 questionnaire*) or (subject-report* near/2 questionnaire*) or (self-administer* near/2 questionnaire*) or (spontaneous* near/2 questionnaire*) or (prompt near/2 questionnaire*) or (prompted near/2 questionnaire*) or (unprompted near/2 questionnaire*) or (open-ended near/2 questionnaire*) or (structured near/2 questionnaire*) or (systematic near/2 questionnaire*) or (standardi* near/2 questionnaire*)):ti,ab

#16 ((elicit*NEAR/2 report?) or (evoke* near/2 report?) or (solicit* near/2 report?) or (unsolicit* near/2 report?) or (self-report? near/2 report?) or (participant-report? near/2 report?) or (subject-report? near/2 report?) or (self-administer* near/2 report?) or (spontaneous* near/2 report?) or (prompt near/2 report?) or (prompted near/2 report?) or (unprompted near/2 report?) or (open-ended near/2 report?) or (structured near/2 report?) or (systematic near/2 report?) or (standardi* near/2 report?)):ti,ab

#17 ((elicit*NEAR/2 enquir*) or (evoke* near/2 enquir*) or (solicit* near/2 enquir*) or (unsolicit* near/2 enquir*) or (self-report* near/2 enquir*) or (participant-report* near/2 enquir*) or (subject-report* near/2 enquir*) or (self-administer* near/2 enquir*) or (spontaneous* near/2 enquir*) or (prompt near/2 enquir*) or (prompted near/2 enquir*) or (unprompted near/2 enquir*) or (open-ended near/2 enquir*) or (structured near/2 enquir*) or (systematic near/2 enquir*) or (standardi* near/2 enquir*)):ti,ab

#18 ((elicit*NEAR/2 inquir*) or (evoke* near/2 inquir*) or (solicit* near/2 inquir*) or (unsolicit* near/2 inquir*) or (self-report* near/2 inquir*) or (participant-report* near/2 inquir*) or (subject-report* near/2 inquir*) or (self-administer* near/2 inquir*) or (spontaneous* near/2 inquir*) or (prompt near/2 inquir*) or (prompted near/2 inquir*) or (unprompted near/2 inquir*) or (open-ended near/2 inquir*) or (structured near/2 inquir*) or (systematic near/2 inquir*) or (standardi* near/2 inquir*)):ti,ab

#19 ((elicit*NEAR/2 checklist*) or (evoke* near/2 checklist*) or (solicit* near/2 checklist*) or (unsolicit* near/2 checklist*) or (self-report* near/2 checklist*) or (participant-report* near/2 checklist*) or (subject-report* near/2 checklist*) or (self-administer* near/2 checklist*) or (spontaneous* near/2 checklist*) or (prompt near/2 checklist*) or (prompted near/2 checklist*) or (unprompted near/2 checklist*) or (open-ended near/2 checklist*) or (structured near/2 checklist*) or (systematic near/2 checklist*) or (standardi* near/2 checklist*)):ti,ab

#20 ((elicit*NEAR/2 check-list*) or (evoke* near/2 check-list*) or (solicit* near/2 check-list*) or (unsolicit* near/2 check-list*) or (self-report* near/2 check-list*) or (participant-report* near/2 check-list*) or (subject-report* near/2 check-list*) or (self-administer* near/2 check-list*) or (spontaneous* near/2 check-list*) or (prompt near/2 check-list*) or (prompted near/2 check-list*) or (unprompted near/2 check-list*) or (open-ended near/2 check-list*) or (structured near/2 check-list*) or (systematic near/2 check-list*) or (standardi* near/2 check-list*)):ti,ab

#21 ((elicit*NEAR/2 query) or (evoke* near/2 query) or (solicit* near/2 query) or (unsolicit* near/2 query) or (self-report* near/2 query) or (participant-report* near/2 query) or (subject-report* near/2 query) or (self-administer* near/2 query) or (spontaneous* near/2 query) or (prompt near/2 query) or (prompted near/2 query) or (unprompted near/2 query) or (open-ended near/2 query) or (structured near/2 query) or (systematic near/2 query) or (standardi* near/2 query)):ti,ab

#22 ((elicit*NEAR/2 querie*) or (evoke* near/2 querie*) or (solicit* near/2 querie*) or (unsolicit* near/2 querie*) or (self-report* near/2 querie*) or (participant-report* near/2 querie*) or (subject-report* near/2 querie*) or (self-administer* near/2 querie*) or (spontaneous* near/2 querie*) or (prompt near/2 querie*) or (prompted near/2 querie*) or (unprompted near/2 querie*) or (open-ended near/2 querie*) or (structured near/2 querie*) or (systematic near/2 querie*) or (standardi* near/2 querie*)):ti,ab

#23 ((elicit*NEAR/2 form?) or (evoke* near/2 form?) or (solicit* near/2 form?) or (unsolicit* near/2 form?) or (self-report* near/2 form?) or (participant-report* near/2 form?) or (subject-

report* near/2 form?) or (self-administer* near/2 form?) or (spontaneous* near/2 form?) or (prompt near/2 form?) or (prompted near/2 form?) or (unprompted near/2 form?) or (open-ended near/2 form?) or (structured near/2 form?) or (systematic near/2 form?) or (standardi* near/2 form?)):ti,ab

#24 ((elicit*NEAR/2 complain*) or (evoke* near/2 complain*) or (solicit* near/2 complain*) or (unsolicit* near/2 complain*) or (self-report* near/2 complain*) or (participant-report* near/2 complain*) or (subject-report* near/2 complain*) or (self-administer* near/2 complain*) or (spontaneous* near/2 complain*) or (prompt near/2 complain*) or (prompted near/2 complain*) or (unprompted near/2 complain*) or (open-ended near/2 complain*) or (structured near/2 complain*) or (systematic near/2 complain*) or (standardi* near/2 complain*)):ti,ab

#25 ("spontaneous report*" or "self report*" or "participant report*" or "patient report*" or "subject report*" or "self administer*"):ti,ab

#26 #15 or #16 or #17 or #18 or #19 or #20 or #21 or #22 or #23 or #24 or #25

#27 #14 and #26

CCTR

Search Name: Elicitation of AE narrower

Last Saved: 02/05/2013 09:50:07.597

Description:

16th March 2015 repeated pubn date 2013-5 retrieved 248

ID Search

#1 MeSH descriptor: [Adverse Drug Reaction Reporting Systems] this term only

#2 MeSH descriptor: [Drug Toxicity] this term only

#3 Any MeSH descriptor with qualifier(s): [Adverse effects - AE, Drug effects - DE]

#4 ((drug? near/2 safety) or (drug? near/2 harm?) or (drug? near/2 adverse) or (drug? near/2 undesirable) or (drug? near/2 tolerability) or (drug? near/2 toxicity) or (drug? near/2 effect?) or (drug? near/2 event?) or (drug? near/2 symptom?)):ti

#5 ((adverse near/2 effect?) or (adverse near/2 reaction?) or (adverse near/2 event?) or (adverse near/2 outcome?) or (adverse near/2 symptom?)):ti

#6 ((side near/2 effect?) or (side near/2 reaction?) or (side near/2 event?) or (side near/2 outcome?) or (side near/2 symptom?)):ti

#7 ((undesirable near/2 effect?) or (undesirable near/2 reaction?) or (undesirable near/2 event?) or (undesirable near/2 outcome?) or (undesirable near/2 symptom?)):ti

#8 (("treatment emergent" near/2 effect?) or ("treatment emergent" near/2 reaction?) or ("treatment emergent" near/2 event?) or ("treatment emergent" near/2 outcome?) or ("treatment emergent" near/2 symptom?)):ti,ab

#9 (("treatment related" near/2 effect?) or ("treatment related" near/2 reaction?) or ("treatment related" near/2 event?) or ("treatment related" near/2 outcome?) or ("treatment related" near/2 symptom?)):ti

#10 ("adverse effect*" near/3 assess*) or ("adverse effect*" near/3 measur*) or ("adverse effect*" near/3 detect*) or ("adverse effect*" near/3 notify) or ("adverse effect*" near/3 notification*):ti,ab

#11 ("adverse event*" near/3 assess*) or ("adverse event *" near/3 measur*) or ("adverse event *" near/3 detect*) or ("adverse event *" near/3 notify) or ("adverse event *" near/3 notification*):ti,ab

#12 ("adverse reaction*" near/3 assess*) or ("adverse reaction *" near/3 measur*) or ("adverse reaction *" near/3 detect*) or ("adverse reaction *" near/3 notify) or ("adverse reaction *" near/3 notification*):ti,ab

#13 ("side effect*" near/3 assess*) or ("side effect*" near/3 measur*) or ("side effect*" near/3 detect*) or ("side effect*" near/3 notify) or ("side effect*" near/3 notification*):ti,ab

#14 #1 or #2 or #3 or #4 or #5 or #6 or #7 or #8 or #9 or #10 or #11 or #12 or #13

#15 ((elicit*NEAR/2 questionnaire*) or (evoke* near/2 questionnaire*) or (solicit* near/2 questionnaire*) or (unsolicit* near/2 questionnaire*) or (self-report* near/2 questionnaire*) or (participant-report* near/2 questionnaire*) or (subject-report* near/2 questionnaire*) or (self-administer* near/2 questionnaire*) or (spontaneous* near/2 questionnaire*) or (prompt near/2 questionnaire*) or (prompted near/2 questionnaire*) or (unprompted near/2 questionnaire*) or (open-ended near/2 questionnaire*) or (structured near/2 questionnaire*) or (systematic near/2 questionnaire*) or (standardi* near/2 questionnaire*)):ti,ab

#16 ((elicit*NEAR/2 report?) or (evoke* near/2 report?) or (solicit* near/2 report?) or (unsolicit* near/2 report?) or (self-report? near/2 report?) or (participant-report? near/2 report?) or (subject-report? near/2 report?) or (self-administer* near/2 report?) or (spontaneous* near/2 report?) or (prompt near/2 report?) or (prompted near/2 report?) or (unprompted near/2 report?) or (open-ended near/2 report?) or (structured near/2 report?) or (systematic near/2 report?) or (standardi* near/2 report?)):ti,ab

#17 ((elicit*NEAR/2 enquir*) or (evoke* near/2 enquir*) or (solicit* near/2 enquir*) or (unsolicit* near/2 enquir*) or (self-report* near/2 enquir*) or (participant-report* near/2 enquir*) or (subject-report* near/2 enquir*) or (self-administer* near/2 enquir*) or (spontaneous* near/2 enquir*) or (prompt near/2 enquir*) or (prompted near/2 enquir*) or (unprompted near/2 enquir*) or (open-ended near/2 enquir*) or (structured near/2 enquir*) or (systematic near/2 enquir*) or (standardi* near/2 enquir*)):ti,ab

#18 ((elicit*NEAR/2 inquir*) or (evoke* near/2 inquir*) or (solicit* near/2 inquir*) or (unsolicit* near/2 inquir*) or (self-report* near/2 inquir*) or (participant-report* near/2 inquir*) or (subject-report* near/2 inquir*) or (self-administer* near/2 inquir*) or (spontaneous* near/2 inquir*) or (prompt near/2 inquir*) or (prompted near/2 inquir*) or (unprompted near/2 inquir*) or (open-

ended near/2 inquir*) or (structured near/2 inquir*) or (systematic near/2 inquir*) or (standardi* near/2 inquir*)):ti,ab

#19 ((elicit*NEAR/2 checklist*) or (evoke* near/2 checklist*) or (solicit* near/2 checklist*) or (unsolicit* near/2 checklist*) or (self-report* near/2 checklist*) or (participant-report* near/2 checklist*) or (subject-report* near/2 checklist*) or (self-administer* near/2 checklist*) or (spontaneous* near/2 checklist*) or (prompt near/2 checklist*) or (prompted near/2 checklist*) or (unprompted near/2 checklist*) or (open-ended near/2 checklist*) or (structured near/2 checklist*) or (systematic near/2 checklist*) or (standardi* near/2 checklist*)):ti,ab

#20 ((elicit*NEAR/2 check-list*) or (evoke* near/2 check-list*) or (solicit* near/2 check-list*) or (unsolicit* near/2 check-list*) or (self-report* near/2 check-list*) or (participant-report* near/2 check-list*) or (subject-report* near/2 check-list*) or (self-administer* near/2 check-list*) or (spontaneous* near/2 check-list*) or (prompt near/2 check-list*) or (prompted near/2 check-list*) or (unprompted near/2 check-list*) or (open-ended near/2 check-list*) or (structured near/2 check-list*) or (systematic near/2 check-list*) or (standardi* near/2 check-list*)):ti,ab

#21 ((elicit*NEAR/2 query) or (evoke* near/2 query) or (solicit* near/2 query) or (unsolicit* near/2 query) or (self-report* near/2 query) or (participant-report* near/2 query) or (subject-report* near/2 query) or (self-administer* near/2 query) or (spontaneous* near/2 query) or (prompt near/2 query) or (prompted near/2 query) or (unprompted near/2 query) or (open-ended near/2 query) or (structured near/2 query) or (systematic near/2 query) or (standardi* near/2 query)):ti,ab

#22 ((elicit*NEAR/2 querie*) or (evoke* near/2 querie*) or (solicit* near/2 querie*) or (unsolicit* near/2 querie*) or (self-report* near/2 querie*) or (participant-report* near/2 querie*) or (subject-report* near/2 querie*) or (self-administer* near/2 querie*) or (spontaneous* near/2 querie*) or (prompt near/2 querie*) or (prompted near/2 querie*) or (unprompted near/2 querie*) or (open-ended near/2 querie*) or (structured near/2 querie*) or (systematic near/2 querie*) or (standardi* near/2 querie*)):ti,ab

#23 ((elicit*NEAR/2 form?) or (evoke* near/2 form?) or (solicit* near/2 form?) or (unsolicit* near/2 form?) or (self-report* near/2 form?) or (participant-report* near/2 form?) or (subject-report* near/2 form?) or (self-administer* near/2 form?) or (spontaneous* near/2 form?) or (prompt near/2 form?) or (prompted near/2 form?) or (unprompted near/2 form?) or (open-ended near/2 form?) or (structured near/2 form?) or (systematic near/2 form?) or (standardi* near/2 form?)):ti,ab

#24 ((elicit*NEAR/2 complain*) or (evoke* near/2 complain*) or (solicit* near/2 complain*) or (unsolicit* near/2 complain*) or (self-report* near/2 complain*) or (participant-report* near/2 complain*) or (subject-report* near/2 complain*) or (self-administer* near/2 complain*) or (spontaneous* near/2 complain*) or (prompt near/2 complain*) or (prompted near/2 complain*) or (unprompted near/2 complain*) or (open-ended near/2 complain*) or (structured near/2 complain*) or (systematic near/2 complain*) or (standardi* near/2 complain*)):ti,ab

#25 ("spontaneous report*" or "self report*" or "participant report*" or "patient report*" or "subject report*" or "self administer*"):ti,ab

#26 #15 or #16 or #17 or #18 or #19 or #20 or #21 or #22 or #23 or #24 or #25

#27 #14 and #26

Supplementary File S4.1: Survey questionnaire (Chapter 4)

Welcome

Elicitation and Recording of Participant-reported Safety Data in Malaria Clinical Drug Trials/Studies

We are interested in talking to malaria clinical research team members about the details of how malaria clinical research participants (or their caregivers, such as a parent) are asked about their health and treatment-taking in order to collect medical history, adverse event and concomitant medication (safety) data. The survey is not about spontaneous reporting of adverse drug reactions (yellow-card, prescription event monitoring etc.) or vaccine clinical research.

This voluntary survey should take about 15-20 minutes. We would like to be able to follow up with you about specific methods you have used and have the opportunity to work with you in future collaborations. We would therefore be grateful if you would provide your name and contact details. These will be stored in secure electronic and paper files and will not be used in any report.
Should other members of your research team have used the same methods to collect safety data then one person may respond on behalf of the team. Ideally, this should be a clinical investigator or someone who has been involved in the selection or development of methods used to collect participant-reported safety data.

1) In which country do you work *most of the time*? ( ) Abkhazia to ( ) Zimbabwe

2) Your name: _____

3) What is the name of the organization where you work? _____

4) Phone number (please include country and area codes!): _____

5) Email address: _____

6) How long have you been involved in *malaria* clinical research?*
( ) Have never been involved in malaria clinical research
( ) < 1 year
( ) 1-5 years
( ) > 5 years

7) Indicate your *most recent primary role* within malaria clinical research? Check the one that fits your role best.*
( ) Working at an investigational site
( ) Representative of a sponsor (other than sponsor-investigator)
( ) Other: _____

Indicate your *most recent primary role* at an investigational site?
( ) Principal investigator
( ) Co-investigator or sub-investigator
( ) Study coordinator

( ) Research nurse
( ) Research team member other (briefly describe role): _____
( ) Other: _____

8) Is the majority of your malaria clinical research sponsored by:
( ) Pharmaceutical companies
( ) Non-commercial entities (e.g. Medicines for Malaria Venture)
( ) Your own institution (i.e. a member of the team is a Sponsor-investigator)
( ) Other: _____

9) Have you had responsibility for *selecting or developing* methods used to collect adverse event or concomitant medication data in clinical trials/studies?
( ) Yes
( ) No

Please consider your *most recent malaria* clinical drug research study where participants (or their caregiver, such as a parent) were asked about their health and use of treatments to collect medical history, adverse event or concomitant medication data.

10) What type of study was your *most recent malaria* clinical research study? Check all that apply:
[ ] Interventional
[ ] Observational
[ ] Single arm
[ ] Multiple arms
[ ] Randomized allocation
[ ] Other

11) What was the research population of your *most recent malaria* clinical research study? Check all that apply:
[ ] Adults
[ ] Children <1year
[ ] Children 1-5 years
[ ] Children 5-12 years
[ ] Children 12-17 years
[ ] Healthy volunteers
[ ] Patients with malaria
[ ] Other

In general, from about what age were children asked directly about their *health?*
( ) Age in years: _____
( ) Children were NOT asked directly at all
( ) Don't know

In general, from about what age were children asked directly about their *use of treatments*?
( ) Age in years: _____
( ) Children were NOT asked directly at all
( ) Don't know

12) Which staff members were involved in asking participants (or their caregivers) about their health and use of treatments? Check all that apply:
[ ] Medical doctor
[ ] Study nurse
[ ] Other

13) Were questions about health and use of treatments ever asked through a translator?
( ) Yes
( ) No
( ) Don't know

Asking participants (or their caregivers) about *health* to collect adverse event (AE) data, at visits *after* the first visit.

14) Were participants (or their caregiver) asked about their *health* since baseline using a general question *without* reference to a particular condition or body system, e.g. 'how have you been feeling?' or 'has your child experienced any problems?'
( ) Yes
( ) No
( ) Don't know

Please give the phrase(s) or question(s) used:
( ) Don't know
( ) Phrase(s) used: _____

Did the study *require* this particular phrase(s) be used for the general question(s)?
( ) Yes
( ) No
( ) Don't know

How confident are you that all staff used the study-specific phrase(s) *most* of the time?
( ) Confident
( ) Not confident

Asking participants (or their caregivers) about their *health* to collect adverse event (AE) data, at visits *after* the first visit (continued).

15) Were participants (or their caregiver) asked about any *change in their health* since baseline using *structured* questions, e.g. 'please tell me if you have experienced fever, cough, headache….?' or 'has your child had any problems with the chest, head….?'
( ) Yes
( ) No
( ) Don't know

Did the study *require particular* questions be asked, e.g. using a prepared list of possible options?
( ) Yes
( ) No
( ) Don't know

How were health issues defined in the questions? Check all that apply:
[ ] By symptom
[ ] By body system
[ ] Expected adverse events
[ ] Malaria symptoms
[ ] Other

How were options presented? Check all that apply:
[ ] Staff used the exact words as prepared
[ ] Staff rephrased items in their own words
[ ] Items were only presented if they had been reported at a previous visit
[ ] Other

Please *describe* any non study-specific *structured* questions asked:
( ) Question(s) used: _____
( ) Don't know

16) Were participants (or their caregiver) asked about their *health* since baseline in *another way* to the general or structured method (e.g. picture tool, diary)?
( ) Yes
( ) No
( ) Don't know

Please briefly describe the method(s): _____

17) What was the rationale for using the above whole approach to questioning about *health* in this study?

18) Was this *same* approach to questioning participants (or their caregivers) about their *health* to collect AEs, also used when asking about medical history at *baseline*?

( ) Yes
( ) No
( ) Don't know

19) How was the questioning about health *different* at baseline? _____

20) Does the above approach to questioning participants (or their caregivers) about their *health* reflect *most malaria clinical studies* you have been involved in?
( ) Yes
( ) No
( ) Don't know
( ) Other: _____

Please briefly describe other questioning methods or combinations of methods used: _____

21) What do you consider as an *optimal* (important and feasible) approach for asking participants (or their caregivers) about their *health* to collect AE data:
( ) The approach reported above
( ) A different approach to my above report (briefly describe): _____

22) Why do you consider that optimal?

Recording of adverse events (AEs) in the database.

23) How were the participants' (or their caregivers') reports eventually *recorded in the database* as AEs? Check all that apply:
[ ] As verbatim participant (or caregiver) reports
[ ] According to staff members' own terminology for symptom or diagnosis
[ ] Standard terminology (e.g. specific coding method)
[ ] Other

Which standard terminology or coding system was used? _____

24) How did you assess AEs for *severity*, if at all?
( ) Not assessed for severity
( ) Published grading scale
( ) Other grading scale method
( ) Don't know

Please provide us with a reference to the grading scale or a description of the method used, if known:

25) How did you assess adverse events for *causality*, if at all?
( ) Not assessed for causality

( ) Published causality rating
( ) Other causality rating method
( ) Don't know

Please provide us with a reference to the rating scale or a description of the method used, if known:

Asking participants (or their caregivers) about their *use of treatments* (other than the study drug) to record previous and concomitant medication data.

26) Were participants (or their caregiver) questioned about their *use of non-study treatments* using a general question *without* reference to a particular treatment class or name, e.g. 'Please tell me about any treatment you have used' or 'Please tell me about any treatment your child is currently using'?
( ) Yes
( ) No
( ) Don't know

Please give the phrase(s) or question(s) used:
( ) Don't know
( ) Phrase(s) used: _____

Did the study *require* this particular phrase(s) be used for the general question(s)?
( ) Yes
( ) No
( ) Don't know

Which of the following were explicitly referred to during general questioning (check all that apply):
[ ] Prescription medicines
[ ] Over the counter medicines
[ ] Traditional treatments
[ ] Supplements
[ ] Vaccinations
[ ] Other

27) Were participants (or their caregiver) questioned about their *use of non-study treatments* by asking *specific* or *structured* questions, e.g. 'Have you taken paracetamol, antibiotics….?'
( ) Yes
( ) No
( ) Don't know

Did the study *require particular questions* be asked, e.g. using a prepared list of possible options?
( ) Yes
( ) No

( ) Don't know

How were non-study treatments defined in these questions? Check all that apply:
[ ] Don't know
[ ] By treatment class (e.g. painkillers)
[ ] By treatment name (e.g. paracetamol)
[ ] Other

Which of the following were explicitly asked about during *structured* questioning? Check all that apply:
[ ] Prescription medicines
[ ] Over the counter medicines
[ ] Traditional treatments
[ ] Supplements
[ ] Vaccinations
[ ] Other

How were options presented? Check all that apply:
[ ] Staff used the exact words as prepared
[ ] Staff re-phrased items in their own words
[ ] Items were only presented if they had been reported at a previous visit
[ ] Other

Please describe any non study-specific questions asked:
( ) Don't know
( ) Question(s) used: _____

28) Were participants (or their caregiver) questioned about their *use of non-study treatments* in *another way* to the general or structured method, e.g. picture tool, diary?
( ) Yes
( ) No
( ) Don't know

Please briefly describe the method(s):

29) What was the rationale for using the above whole approach for questioning about *use of non-study treatments* in this study?

30) Does the above approach to questioning participants (or their caregivers) about their *use of non-study treatments* reflect normal practice in *most* malaria clinical studies you have been involved in?
( ) Yes
( ) No
( ) Other: _____

Please briefly describe other questioning methods or combinations of methods used:

31) What do you consider as an *optimal* (important and feasible) approach for asking participants (or caregivers) about their *use of non-study treatments* to collect previous and concomitant medication data:
( ) The approach reported above
( ) A different approach to my above report (briefly describe): _____

32) Why do you consider that optimal?
Please consider your *most recent malaria* clinical research study where participants (or their caregiver, such as a parent) were asked about adherence to treatment(s) given during the study, when answering the questions on this page.

33) What was the study drug regimen?

34) Did you collect individual participant data on drug intake?
( ) Yes
( ) No
( ) Don't know

35) What was the reason for not collecting such data?
( ) Don't know
( ) Reason(s): _____

36) How did you collect the data? Check all that apply:
[ ] Dose observed
[ ] Participant (or caregiver) recall
[ ] Pill count (manual/electronic)
[ ] Dispensing confirmation
[ ] Pill diary
[ ] Other

Were all drug doses observed?
( ) All doses observed
( ) Some doses observed
( ) Don't know

37) Which data did you record? Check all that apply:
[ ] Quantity of dose dispensed
[ ] Dose given
[ ] Duration of total therapy
[ ] Time of dose

[ ] Whether participant vomited
[ ] Reason(s) for non-adherence
[ ] Whether taken with food/drink
[ ] Other

Was time of dose recorded for all drug doses?
( ) Time was recorded for all doses
( ) Time was recorded for some doses
( ) Don't know

When it was recorded that a participant vomited, what action was taken and how was this recorded?

38) How did you define adherence levels in your study?

39) Did you report adherence in the trial report?
( ) Yes
( ) No
( ) Don't know

General

41) We would appreciate if you would suggest any *references* that you know to be relevant for this work. Please describe any references you may know of here:

42) Would you like to hear about future projects for this topic?
[ ] Yes

Thank you for taking the time to complete this survey, we appreciate your input into this complex topic.

Supplementary File S5.1: Participant in-depth interview questionnaire (Chapter 5)

Participant ID: |IDI|__|__|__|__|__|__|* Facilitator's initials: |__|__|__| Co-facilitator's initials: |__|__|__|
* IDI, 3 digit study code (SEA=SEACAT, INT=INTERACT), sequential interview in the study site (e.g. IDISEA01)

Date (ddmmmyy):|_|_|_|_|_|_|_|

Introduction (including the following points to be covered)
□ I am not part of the clinical study team although I do work in the same department and I am familiar with this study. I am not a doctor, I am trained in talking to people about their lives and I am particularly interested in how we find out whether treatments work and are safe.
□ I am very interested to talk to you about how you became involved with this study, your health and what you take to keep healthy. I expect this to take about an hour.
□ I will record our conversation on a tape recorder and the words we both speak will be typed into a computer. This will help me to remember what we spoke about and I can then compare your experiences with those of other people I speak to.
□ There is no right or wrong answer to any of my questions, just your experiences, opinions and ideas which are all valuable; it's important for me to hear all sides of an issue – the positive and negative – as this will help us to make these studies the best that we can.
□ Your confidentiality is important to me. As for all information you tell us about this study, you will never be identified by your name in any report or to anyone outside of this team.
□ Do you have any questions?
□ Consent/assent

Warm up
Tell me a bit about yourself; where do you live, how long have you lived in this community, do you work, who do you live with?

Part I health status and treatments taken prior to and during study (probe all references to health/treatments in terms of detail including dates).

Introductory
How did you come to hear about this study? *(Probes: which clinics do you normallyattend, doctors/other healthcare workers do you see?)*

Narrative
Please tell me what happened in this trial from the first time you saw Dr X *(Probes: What happened next? What did the trial staff do? What did they ask you about? If health-related: what was your health like? If treatment-related: what treatments were you taking?)*

Wait for full response and take notes of any relevant illness/treatments either part of medical/drug history or treatment-emergent/concomitant medications so these may be re-visited in next sections

Medical/drug history
Let's go back to before the trial. You first saw the Dr in [give date], can you tell me [more] about your health then? *(Probes: What was that like? What did you do about that? Who did you speak to about that? What did they say?)*
And the month before that [give month]? *(Probes: As above)*
Have you had any [other] times in your life when you were quite ill? *(Probes: As above for previous malaria and illness other than malaria)*
Were there issues with your health when you were not as ill? *(Probes: As above)*

What health issues are common for people with malaria? *(Probes: did you have those,if so, when, what did you do about that? Who did you speak to about that? What did they say?)*

When is it necessary to have treatment for malaria? *(Probes: when should you start treatment? What is the best treatment to start with? When is no treatment required? Have you taken those for this malaria, if so, what/when?)*

[For those with HIV] What health issues do you think are common for people with HIV? *(Probes: have you had any of those, if so, when? When is it necessary to have treatment, when is no treatment required? Do you feel that this can be prevented, if so, how? Have you taken those, if so, when?*)
Are there some [other] things that people with HIV take normally for their health? *(Probes: Why is that? For general health, for illnesses, prescription medicine, over-the-counter, traditional/complementary remedies? Have you taken those, if so, when?)*

Health in trial
Can you tell me [more] about your health since you joined the trial? *(Probes: did you notice any changes to your health? How did it change?)*
And since you have been at home *(Probes: As above)*
Probe each symptom reported in the interview so far that was treatment-emergent:*What did you do about that? Who did you speak to about that? What was that like for you?*
How would you rate that in terms of how ill, or not so ill, you were? *What makes you say it was [mild/moderate/severe- equivalents]? What do you think was the reason for that change in your health?)*

Drugs in trial
Did you take/use anything [else] for any other reason? *(Probe: For example, did you take anything else to help with the malaria/prevent ill health/just for general good health/contraception/vitamins, traditional/alternative medicines?)*
Who do you normally see for your health issues? *(Probe: some people [also] speak to their family/friends/a pharmacist/shopkeeper/a sangoma/alternative health practitioner)*
When was the last time you spoke to any of those people about your health? *(Probe: what brought you to see that person? What did they recommend?)*

Show example products and discuss if taken/given

Part II reasons for differences in reporting depending on the method and help in accessing all information (probe all references to health/treatments)

Reporting
So, let's look at your file *(get file from within a pile on the shelf)*. I'd really like your help finding the best way for the clinic to collect this information so I'd like to see what they wrote down and how it compares with what we've been discussing.

It does not matter if we find there is a difference between what the clinic wrote down and what we have discussed today – there is no right or wrong and you will not be in any trouble either way

Browse completed forms together highlighting differences between the general enquiry responses, the checklist responses and any additional previously unreported Med history, AEs, conc meds you have noted during the interview

If discrepancy between data collection forms and also with the interview:
Why do you think the clinic didn't get that information that time? *(Probes: What helps you remember some things better than others? Were there any things that the clinic asked that you weren't sure about? What makes you feel that you needed to/didn't need to tell the clinic about x? What do you think about the first question (give example of open enquiries) and the checklists? And the way we are talking today?)*

For each answer try to probe further e.g. if he/she says something was not important to report ask why, and which things are important to report/which are not)

For AEs in clinic file not discussed today – probe each symptom for severity as above

Do you think what we have discussed captures your health and treatments? What was it like for you to talk about your health and what you have taken/used/given with the doctors and nurses for this trial?
If we want to get the most information on the health and treatments of people who take part in studies like this, how would you suggest we do this? *(Probes: diaries, pictures of medicines, who should ask them the questions (doctor, nurse, community health worker, traditional healer), what type of questions)*

We are now coming to the end of our talk. Is there anything else anyone would like to add about the study, your health, treatments and how we can get the most information?

□ Summarise □ Thank participant

Supplementary File S5.2: Participant focus group discussion questionnaire (Chapter 5)

Focus group ID: |FG|__|__|__|__|__|* Facilitator's initials: |__|__|__| Co-facilitator's initials: |__|__|__|
* FG, 3 digit study code (SEA=SEACAT, INT=INTERACT), participant group (P=patient, S=staff), sequential group at study site (e.g. FGSEAP01)
Date (ddmmmyy): |_|_|_|_|_|_|_|

Introduction (including the following points to be covered)
□ Introduce facilitator/co-facilitator and roles. We are not part of the clinical trial team although I do work in the same department/I work for _____ and I am familiar with this trial. We are not doctors but are trained in talking to people about their lives and are particularly interested in how we find out whether treatments work and are safe.
□ We are very interested to talk to you about what treatments you and other people in your community take for illness and to keep healthy and what was your experience of the trial. We hope that you will talk to each other about the things that I mention over the next hour or so.
□ We will record our conversation on a tape recorder and the words we both speak will be typed into a computer. This will help me to remember what we spoke about and I can then compare your experiences with those of other people I speak to.
□ Please could we have one person speaking at a time; it is important to hear everyone's ideas and opinions. There is no right or wrong answers, just your experiences, opinions and ideas which are all valuable; it's important for me to hear all sides of an issue – the positive and negative – as this will help us to make these studies the best that we can.
□ Your confidentiality is important to us and I would like you to respect each other's confidentiality too. Whatever is said within this room should remain in the room! You will never be identified by your name in any report or to anyone outside of this team. In fact, whatever you tell me now cannot be linked to your participation in the clinical trial as we use different identification codes so you are anonymous. You can be as free as you like with the information you tell us.
□ Do you have any questions?
□ Consent

General: make a separate note of anyone mentioned specific illness/treatment during the trial and ask these people at the end of section on barriers to reporting specifically if they reported this information in the trial and why/why not.

Warm up
Tell us a little bit about yourselves; where do you live, how long have you lived in this community?

Introductory
Please remind me which clinics do you normally attend?

Perceptions of illness & common treatment practices
Let's talk about malaria. *W*hat conditions/symptoms are common for people who get malaria? *(Probes: What do you do about that? Who do you speak to for advice and treatment – first, and after that?)*
[HIV group only] What about people with HIV who get malaria? *(Probes: Is it different? What do you do about that? Who do you speak to for advice and treatment –first, and after that?)*
So when is it necessary to have treatment for malaria? *(Probes: when should you start treatment? What is the best treatment to start with? When is it necessary to see a healthcare worker? When is no treatment required? Have you taken those, if so, what/when?)* Try to ascertain if it was within timescale of trial.
[HIV group only] What about people with HIV who get malaria? *(Probes: is it different? When should you start treatment? What is the best treatment to start with? When is no treatment required? Have you taken those, if so, what/when?)*

What other things do people with malaria use normally? *(Probes: for general health for illnesses, cleansing/local traditions? Which of these have any of you taken? Why did you take it?)* Try to ascertain if it was within timescale of trial.
[HIV group only] What other things do people with HIV use normally? *(Probes: for general health, for illnesses, cleansing/local traditions? Which of these have any of you taken? Why did you take it?)* Try to ascertain if it was within timescale of trial.
[HIV group only] Do people who are taking ARVs need different treatments to people who are not taking ARVs? *(Probes: why is that? for general health, for illnesses? Which of these have any of you taken? Why did you take it? When did you take it?)*
Who do you normally see for your health issues? *(Probe: some people [also] speak to their [family/friends] [a pharmacist] [shopkeeper] [sangoma]* What illnesses can you deal with without seeing a healthcare worker?*)*

Did anyone see these people during the trial? *(Probe: why did you see that person, what did they recommend for you?)*

[HIV ARV group only] Does it depend on whether someone is on ARVs or not who you choose to see for health issues? *(Probes: Why is that?)*
Does being in a trial change who you see for health issues? *(Probes: Why is that?)*
And does it change the treatments you use? *(Probes: Why is that?)*

Experience of trial
Would somebody like to tell me about the most important thing that happened to them in the trial? *(Probe: Wait for full response before probing if relevant to health/treatment:*
Can I take you back to…. How were you feeling at that time? What did you take for that? When did you talk to the trial staff about that?)
Anybody else? *(Probe: did you have a different important experience?)*

Barriers to reporting & experience of AEs
How did the trial staff ask you about your health when you started the trial? *(Probe: Who asked? What did they ask? What information were they looking for? Why do you think they were asking you about that? How did you feel about being asked?)*
And about what you were taking at the start of the trial? *(Probe: As above)*
How did the trial staff ask you about your ongoing health during the trial? *(Probe: As above)*
And about what you were taking? *(as above)*
What did you think about the different questions about your health and treatments [give example of open enquiry and checklist?
Are there some things the trial staff would not need to know? (*Probe: Which illnesses/treatments do you think are not important enough to mention?*)
Are there some things the trial staff would not want to know? (*Probe: Which illnesses/treatments do you think are not important enough to mention?*)
Which health issues, if any, are easier to tell the trial staff about? *(Probe: severity, personal, caused by medicine?)*
Which treatments, if any, are easier to tell the trial staff about? *(Probe: is it different for ARVs, prescriptions, over-the-counter, vitamins, traditional meds,* how often used*?)*
Which times in the trial, if any, are better for speaking about health and treatment?
Which places in the trial, if any, are better for speaking about health and treatments?
Which staff in the trial, if any, is better for speaking to about health and treatments?
Did anyone have a change in health during the trial that they didn't tell the trial staff about? *(Probe: What was that? Why was that?)*
Did anyone take anything during the trial that they didn't tell the trial staff about? *(Probe: What was that? Why was that?)*
Do you think the trial staff got everything what happened with your health and treatments adequately? *(Probe: Why did it/didn't it?)*

Some people worry about mentioning an illness because they think that they may be doing something wrong or asked to leave the trial. What do you think about that?

Help with reporting

If we want to get the most information on the health and treatments of people who take part in studies like this, how would you suggest we do this? *(Probe: What helps you remember some things better than others? Were there any things that the clinic asked about that you weren't sure about? What makes you feel that you do/do not need to tell the clinic about some things? Were you afraid of giving some information? Did they use any words that you didn't understand?)*

Closing

We are now coming to the end of our talk. Is there anything else anyone would like to add about the study, your health, treatments and how we can get the most information in clinical trials?

□ Summarise
□ Thank participants
□ Provide extra information and contacts to participants
□ Collect participant demographic details

Supplementary File S5.3: Staff focus group discussion (Chapter 5)

Focus group ID: |FG|__|__|__|__|__|__|* Facilitator's initials: |__|__|__| Co-facilitator's initials: |__|__|__|
* FG, 3 digit study code (SEA=SEACAT, INT=INTERACT), participant group (P=patient, S=staff), sequential group at study site (e.g. FGSEAS01)
Date (ddmmmyy): |_|_|_|_|_|_|_|_|

Introduction (including the following points to be covered)
□ Introduce facilitator/co-facilitator and roles.
□ We are very interested to talk to you about what methods are used to elicit information about trial participants' medical and drug histories, adverse events (AEs) and concomitant medications. We hope that you will talk to each other about the issues that I raise over the next hour or so.
□ We will be recording our conversation on a tape recorder and a transcript will be entered into a computer program. This helps us to remember what we spoke about and I can then compare your experiences with those of other people I speak to.
□ Please could we have one person speaking at a time; it is important to hear everyone's ideas and opinions. There is no right or wrong answers, just your experiences, opinions and ideas which are all valuable; it's important for me to hear all sides of an issue – the positive and negative – as this will help us to make these studies the best that we can.
□ Your confidentiality is important and I would like you to respect each other's confidentiality too. Whatever is said within this room should remain in the room! You will never be identified by your name in any report or to anyone outside of this team. You can be as free as you like with the information you tell us.
□ Do you have any questions?
□ Consent

Warm up
Tell me a little bit about your role in the trial and how many other trials you have worked on where you questioned trial participants about their medical and drug information.

Domain Topic and probes

Introductory
What do you feel we need to know about trial participants' medical history? *(Probes: What extent of data is necessary? Are there particular conditions that are more important than others to know? Why do we need these data?)*
And the treatments they have taken before the trial? *(Probes: What extent of data is necessary Are there particular treatments that are more important than others to know? Why do we need these data?)*
What about their health during the trial? *(Probes: What extent of data is necessary Are there particular conditions that are more important than others to know? Why do we need these data?)*
And the treatments they take during the trial? *(Probes: What extent of data is necessary? Are there particular treatments that are more important than others to? Why do we need these data?)*

Elicitation methods
Let's talk about how we ask trial participants about their medical history. Would anyone like to describe a method you are familiar with? *(Probes: what do you feel about this method? Does it work well? Why is that?)*
What is most important to you when asking patients about their medical history?
Anyone else have experience of the same or a different method? *(Probe as above and repeat until all types of methods have been discussed)*
Do you think one method is better than the others? *(Probe: why is that?)*
Should we use a combination of methods? *(Probe: why is that?)*

What about asking participants about their treatment history? Would anyone like to describe a method you are familiar with? *(Probes as above)*
What is most important to you when asking patients about their treatment history?

What about asking participants about adverse events? *(Probes as above)*
What about asking participants about concomitant medication? *(Probes as above)*

Barriers to reporting
Do you know or believe that participants do/do not always tell you everything when you use x method to obtain medical history data? *(Probes: which data do they tell/don't tell? Why do you think? What are the implications of under-reporting?)*
Does anyone have an(other) example of when you find out that a participant has a certain medical history but this was not offered at the study start? (*Probe: how did you handle that? Why do you think that happened? Does this matter - what are the implications of that?)*
What about y method? *(Probe as above and continue for all methods for all data type. For drugs - what about prescription medicines, over-the-counter, traditional medicines, vitamins, contraceptives?)*

Optimal methods
If we want to get the most information on the health and treatments of people who take part in studies like this, how would you suggest we do this? *(Probes: diaries, pictures of medicines, who should ask them the questions (doctor, nurse, community health worker, traditional healer), what type of questions)*
Does anybody else have a different suggestion they would like to share? *(Probe as above for all data types?)*

Closing
We are now coming to the end of our talk. Is there anything else anyone would like to add about the study, your health, treatments and how we can get the most information in clinical trials?

□ Summarise
□ Thank participants
□ Provide extra information and contacts to participants
□ Collect participant demographic details

Supplementary File S6.1: Participant focus group discussion questionnaire (Chapter 6)

Any changes to the questions should be coordinated through the international pregnancy registry pilot study team as there are implications with regard to version control and eventual pooling of data.

FGD ID: |FG|__|__|__|__|__|__|* Facilitator's initials: |__|__|__| Co-facilitator's initials: |__|__|__|
* FG, 2 digit country code, 2 digit site code, sequential number in the country (3 digits e.g. GHDO03)

Participant sub-group: Registry □ Community □ Date (ddmmmyy):
|_|_|_|_|_|_|_|_|

Introduction
□ Introduce facilitator/co-facilitator
□ Aims of the FGD and expected duration: The investigators of this study would like to collect information on what women in this community think about pregnancy and what they do for health and treatments during pregnancy. This information will be helpful for improving our care to pregnant women. We hope that you will talk to each other about the things that I mention over the next hour and a half.
□ It is important that everyone contributes to the conversation as we are interested in all your experiences.
□ What will happen to the collected information and how the participant/target group may benefit: we will be recording our conversation on a tape recorder and the words spoken will be typed into a computer. This will help us to compare your experiences with those of women in other places.
□ Ground rules: one person at a time; importance of hearing everyone's ideas and opinions, no right or wrong answers to questions just ideas, experiences and opinions which are all valuable; important to hear all sides of an issue – the positive and negative, confidentiality is important "what is shared in the room stays in the room". Any further rules from the group?
□ Any questions?
□ Consent

Warm up
Can we go around the room and say your name, how long each of you has lived in this place, the number of children you have and how many weeks/months pregnant you are.

Introductory (keep short!)
I'd like us to talk about pregnancy. How did you know that you were pregnant?(Probes: symptoms (incl. missed period), test (by whom), when did you know, did you keep a record of periods/date of pregnancy?
Who did you tell about your pregnancy? Probes: when did you consult a healthcare worker? [if delay] what were the reasons for not consulting a healthcare worker earlier? What would it take for you to seek ANC care earlier?

Treatment-seeking/reporting
Let's talk about health. Have any of you felt unwell during this or another pregnancy? *Probes: what did you do (who did you speak to, what did they advise you to take/do)? Can anyone else describe a health problem they or someone they know had?*
What conditions are normal or common in pregnancy? *Probes: has anyone experienced these (or not: can anyone tell us about someone else they know who experienced these?) How was this treated?*
What about [other] treatments during pregnancy. What [other] treatments have you taken during any pregnancy? *Probes: was this prescription medicine, over-the-counter, traditional remedy? What did you take that for? Can anyone else describe a treatment they (or someone they know) take?*
Which traditional remedies/herbs are taken by pregnant women? *Probes: what for? When? Have any of you experience of taking these?*

When is it necessary to have treatment in pregnancy? *Probes: why is that? What is taken?*
Are there [other] things that pregnant women should be taking? *Probes: for the general health of the mother/baby, for illnesses? What? Why?*

[Optional]
I am interested in malaria. How is malaria in pregnancy treated? *Probes: prescription medicine, over-the-counter, traditional remedies, what are the reasons for not consulting a particular person?*
How is malaria in pregnancy prevented? *Probes: prescription medicine, over-the-counter, traditional remedies, what are the reasons for not consulting a particular person?*
What about other conditions such as HIV? How is HIV in pregnancy dealt with? *Probes: prescription medicine, over-the-counter, traditional remedies, what are the reasons for not consulting a particular person?*

How is HIV in the baby prevented? *Probes: prescription medicine, over-the-counter, traditional remedies, what are the reasons for not consulting a particular person?*
Who do you speak to about health and treatments during your pregnancies? *Probes: partner/family/friends/healthcare worker, what are the reasons for telling/not telling these people, do you keep a record of these consultations?*
What should not be taken by pregnant women? *Probes: why should these not be taken?*
Does anyone keep a record of consultations or treatments? *Probes: how? Why?*

Perception of poor birth outcomes CARE – use sensitivity e.g. "this question is not related to you and your baby"
I'd now like for us to talk about babies who are born with problems. Do you know of babies in your community who were born with problems (give examples of pregnancy loss & birth defects, if necessary [for birth defects, start with mild e.g. extra finger, then more severe, e.g. missing limb])?
Why do you think this happens? *Probes: Is it because of some medical illness in the mother/father? Something she/he ate or took or did? Are some things taken/eaten/done to prevent or treat problems in unborn babies/keep the mother and unborn baby healthy?*
[For birth defects] How are such babies treated by family members and members of the community? *Probes: Are they treated and cared for the same as other children? If they experience health problems, how and where are they treated? Do they have the same opportunities as other children e.g. going to school, play etc.?*
If a child is found to have a serious problem at birth, how should the parents be told? *Probes: Who should be the one to tell (health worker, community member, midwife, religious leader, government)? Who should parents be with when they get the news? What kind of information would parents want to know about the problem?*

Pregnancy registry (enrolled)
Let's discuss the pregnancy registry. What do you think about your experience of being part of the registry? *Probes: is it useful/not useful, necessary/not necessary*
Would you advise a friend or relative to take part in this registry? *Probes: why/why not?*
I'd like us to go back to the things we discussed earlier about treatments during pregnancy (review treatments discussed). Which of these need be told to an ANC nurse when she asks about treatments in pregnancy? *Probes: What makes you feel that these need to/don't need to be told? What other things need to be/don't need to be told to the nurse?*
What are the reasons why women may not tell an ANC nurse about things they have taken during a pregnancy? *Probes: do they remember everything?*
Did you not tell the ANC nurse about treatments taken in your pregnancy in this registry? *Probe: why was that? What can we do to help you tell the nurse about all treatments? What about alcohol use?*
If we want to get the most information on what treatments have been taken, how would you suggest we do this? *Probes: diaries, pictures of medicines, who should ask them the questions (doctor, nurse, community health worker, traditional healer), what type of questions?*

Pregnancy registry (community)

Let's discuss the pregnancy registry *("A pregnancy registry is a study where pregnant registry women are invited to be followed up throughout their pregnancy until the time when they deliver their baby. Information on the mother's health and treatments she has taken are recorded and studied, together with information on the health of the baby at the time of its birth*").

What do you think about the idea of being part of the registry? *Probes: is it useful/not useful, necessary/not necessary? Why would you take part/not take part?*
I'd like us to go back to the things we discussed earlier about treatments during pregnancy (review treatments discussed). Which of these need be told to an ANC nurse if she asks about treatments in pregnancy? *Probes: What makes you feel that these need to/don't need to be told? What other things need to be/don't need to be told to the nurse?*
What are the reasons why women may not tell an ANC nurse about things they have taken during a pregnancy? *Probes: do they remember everything?*
Have there been times in you previous pregnancy/ies when you didn't tell the ANC nurse about treatments taken? *Probe: why was that? What can we do to help you tell the nurse about all treatments? What about alcohol use?*
If we want to get the most information on what treatments have been taken, how would you suggest we do this? *Probes: diaries, pictures of medicines, who should ask them the questions (doctor, nurse, community health worker, traditional healer), what type of questions?*

Closing

We are now approaching the end of our discussion. Is there anything else anyone would like to add about the pregnancy registry, treatments pregnant women take/report and babies born with problems?

□ Summarise
□ Thank participants
□ Provide extra information and contacts to participants
□ Collect participant demographic details

UNIVERSITY OF CAPE TOWN

Health Sciences Faculty
Research Ethics Committee
Room E52-24 Groote Schuur Hospital Old Main Building
Observatory 7925
Telephone [021] 406 6338 • Facsimile [021] 406 6411
e-mail: sumayah.ariefdien@uct.ac.za

23 October 2009

REC REF: 376/2009

Prof Barnes
Pharmacology

Dear Prof Barnes

PROJECT TITLE: OPTIMISING THE ACCURATE ELICITATION OF PARTICIPANT-REPORTED DATA RELATED TO SAFETY.

Thank you for your comprehensive response.

It is a pleasure to inform you that the Ethics Committee has **formally approved** the above-mentioned study.

Approval is granted for one year till the 30th October 2010.

Please submit an annual progress report if the research continues beyond the expiry date. Please submit a brief summary of findings if you complete the study within the approval period so that we can close our file.

The following documentation is approved:-

1. Protocol SEACAT 2.5.1 optimising the accurate elicitation of participant--reported data related to safety, dated 19 October 2009.
2. Appendix 1-IDI question guide. SEACAT 2.5.1. App1_IDI-question guide-191009.
3. Appendix 2- FGD question guide: patients SEACAT 2.5.1_App2_FGD question guide-patients_191009
4. Appendix 4- Informed Consent Form for in-depth interview. SEACAT 2.5.1-App 4 –IDI_191009.
5. Appendix 5-Informed Consent Form for FGD dated 191009
6. Appendix6-Informed Consent Form for FGD for staff, dated 310809
7. Appendix 7-Assent Form for in-depth interview, dated 310809

Please note that the ongoing ethical conduct of the study remains the responsibility of the principal investigator.

Please quote the REC. REF in all your correspondence.

ariefdien

Yours sincerely

PROFESSOR M BLOCKMAN
CHAIRPERSON, HSF HUMAN ETHICS

Federal Wide Assurance Number: FWA00001637.
Institutional Review Board (IRB) number: IRB00001938

This serves to confirm that the University of Cape Town Research Ethics Committee complies to the Ethics Standards for Clinical Research with a new drug in patients, based on the Medical Research Council (MRC-SA), Food and Drug Administration (FDA-USA), International Convention on Harmonisation Good Clinical Practice (ICH GCP) and Declaration of Helsinki guidelines.

The Research Ethics Committee granting this approval is in compliance with the ICH Harmonised Tripartite Guidelines E6: Note for Guidance on Good Clinical Practice (CPMP/ICH/135/95) and FDA Code Federal Regulation Part 50, 56 and 312.

ariefdien

THE UNITED REPUBLIC OF
TANZANIA

National Institute for Medical Research
P.O. Box 9653
Dar es Salaam
Tel. 255 22 2121-900/390
Fax. 255 22 2121380/2121360
E-mail: headquarters@nimr.or.tz

NIMR/HQ/R.8a/Vol. IX/1903

Ministry of Health and Social Welfare
P.O. Box 9083
Dar es Salaam
Tel: 255 22 2120262-7
Fax: 255 22 2110986

15th September 2010

Elizabeth Allen
University of Cape Town
Division of Clinical Pharmacology
K45 Old Main Building, Groote Schuur Hospital
Observatory 7925, South Africa

CLEARANCE CERTIFICATE FOR CONDUCTING MEDICAL RESEARCH IN TANZANIA

This is to certify that the research entitled: Protocol SAMCP 2.5.1 Optimizing the Accurate Elicitation of Participant-Reported Data Related to Safety, J Allen E *et al*), whose Local Investigator is Dr Martha Lemnge, NIMR, Tanga, has been granted ethics clearance to be conducted in Tanzania.
The Principal Investigator in the study must ensure that the following conditions are fulfilled:

1. Annual Progress report is to be submitted to the Ministry of Health and the National Institute for Medical Research, Regional and District Medical Officers.
2. Permission to publish the results obtained from National Institute for Medical Research.
3. Copies of final publications are made available to the Ministry of Health & Social Welfare and the National Institute for Medical Research.
4. Any researcher who contravenes or fails to comply with these conditions, shall be guilty of an offence and shall be liable on conviction to a Fine, as MR Act No. 23 of 1979, PART II Section 10(2).
5. Approval is for one year, 13th September 2010 to 12th September 2011

Name: Dr Mwelecele N Malecela

Signature
ACTING CHAIRPERSON
MEDICAL RESEARCH
COORDINATING COMMITTEE

Name: Dr Deo M Mtasiwa

Signature
CHIEF MEDICAL OFFICER
MINISTRY OF HEALTH, SOCIAL WELFARE

CC: RMO
DMO

GHANA HEALTH SERVICE ETHICAL REVIEW COMMITTEE

In case of reply the number and date of this Letter should be quoted

Research & Development Division
Ghana Health Service
P. O. Box MB 190
Accra

My Ref: GHS-ERC 3
Our Ref. No.

23rd December 2009

Tel: +233-21-681106
Fax: + 233-21-229729
Email: Hannah.Frimming@hrw-ghs.org

DR. CHRITINE A. CLERK, Principal Investigator

ETHICAL CLEARANCE – ID NO: GHS-ERC 02/3/09

The Ghana Health Service Ethics Review Committee has reviewed and given approval for the implementation of your Study Protocol titled:

"A PILOT STUDY TO ASSESS THE FEASIBILITY OF A MEDICINE IN PREGNANCY REGISTRY IN GHANA"

This approval requires that you submit periodic review of the protocol to the Committee and a final full review to the Ethical Review Committee (ERC) on completion of the study. The ERC may observe or advise to be observed procedures and records of the study during and after implementation.

Please note that any modification of the project must be submitted to the ERC for review and approval before its implementation.

You are also required to report all serious adverse events related to this study to the ERC within seven days verbally and fourteen days in writing.

You are requested to submit a final report on the study to assure the ERC that the project was implemented as per approved protocol. You are also to inform the ERC and your mother organization before any publication of the research findings.

Please always quote the protocol identification number in all future correspondence in relation to this protocol.

SIGNED.....
PROFESSOR ALBERT GEORGE BAIDOE AMOAH
(GHS-ERC CHAIRMAN)

Cc: The Director, Research & Development Division, Ghana Health Service, Accra

MAKERERE UNIVERSITY
P. O. Box 7072 Kampala Uganda
E-mail: dean@med.mak.ac.ug
Phone: 256 414 530020
Fax: 256 414 541036/0414 532204

FACULTY OF MEDICINE
OFFICE OF THE DEAN

January 20, 2010

Dr. Josephat Byamugisha
Department of Obstetrics and Gynaecology

Dear Dr. Byamugisha,

Re: **Approval of Protocol #REC REF 2010- 013**
"Pilot study to assess the feasibility of medicines in pregnancy registry"

Thank you for submitting an application for approval of the above – referenced protocol. The committee reviewed it and granted approval for one year, effective January 20, 2010. Approval will expire on January 19, 2011.

Continuing Review
In order to continue work on this study (including data analysis) beyond the expiration date, the Faculty of Medicine Research and Ethics Committee must reapprove the protocol after conducting a substantive, meaningful, continuing review.
This means that you must submit a continuing report form as a request for continuing review. To best avoid a lapse, you should submit the request six (6) to eight (8) weeks before the lapse date. Please use the forms supplied by our office.

Amendments
During the approval period, if you propose any change to the protocol such as its funding source, recruiting materials, or consent documents, you must seek Faculty of Medicine Research and Ethics Committee approval before implementing it.
Please summarize the proposed change and the rationale for it in a letter to the Faculty of Medicine Research and Ethics Committee. In addition, submit three (3) copies of an updated version of your original protocol application- one showing all proposed changes in bold or 'track changes,' and the other without bold or track changes.

Reporting
Other events which must be reported promptly in writing to the Faculty of Medicine Research and Ethics Committee include:
Suspension or termination of the protocol by you or the grantor

Unexpected problems involving risk to participants or others

Adverse events, including unanticipated or anticipated but severe physical harm to participants.

Do not hesitate to contact us if you have any questions. Thank you for your cooperation and commitment to the protection of human subjects in research.

Final approval is to be granted by Uganda National Council of Science and Technology.

Yours sincerely,

MAKERERE UNIVERSITY
FACULTY OF MEDICINE
★ 20 JAN 2010 ★
RESEARCH & ETHICS COMMITTEE
P. O. BOX 7072, KAMPALA

Dr. Charles Ebinga
Chairperson Faculty of Medicine Research and Ethics Committee

INSTITUTIONAL RESEARCH AND ETHICS COMMITTEE (IREC)

MOI TEACHING AND REFERRAL HOSPITAL
P.O. BOX 3
ELDORET
Tel: 33471/2/3

MOI UNIVERSITY
SCHOOL OF MEDICINE
P.O. BOX 4606
ELDORET
Tel: 33471/2/3

Reference: IREC/2009/90

11th May, 2010

Prof. E. Were & Team,
Moi University,
School of Medicine,
P.O. Box 4606-30100,
ELDORET.

Dear Prof. Were & Team,

RE: EXPEDITED APPROVAL

The Institutional Research and Ethics Committee has reviewed your proposal titled:

"A Pilot Study to Assess the Feasibility of in a Medicines in Pregnancy Registry Bungoma District".

The proposal has been granted an Expedited Approval on 11th May, 2010 and you are therefore permitted to commence your investigations.

Note that this approval is for 1 year; it will thus expire on 10th May, 2011. If it is necessary to continue with this research beyond the expiry date, a request for continuation should be made in writing to IREC Secretariat two months prior to the expiry date.

You are required to submit progress report(s) regularly as dictated by your proposal. Furthermore, you must notify the Committee of any proposal change (s) or amendment (s), serious or unexpected outcomes related to the conduct of the study, or study termination for any reason. The Committee expects to receive a final report at the end of the study.

Yours Sincerely,

INSTITUTIONAL RESEARCH & ETHICS COMMITTEE
APPROVED
11 MAY 2010
P.O. Box 4606 ELDORET

DR. OMAR [illegible]
CHAIRMAN
INSTITUTIONAL RESEARCH AND ETHICS COMMITTEE

cc: Director - MTRH
Dean - SOM
Dean - SPH
Dean - SOD

UNIVERSITY OF CAPE TOWN
Faculty of Health Sciences
Human Research Ethics Committee

Room E52-24 Old Main Building
Groote Schuur Hospital
Observatory 7925
Telephone [021] 406 6338 • **Facsimile** [021] 406 6411
Email: lornees.emjedi@uct.ac.za
Website: www.health.uct.ac.za/research/humanethics/forms

07 January 2014

HREC REF: 760/2013

Mrs E Allen
Medicine/Clinical Pharmacology
K Floor
OMB

Dear Mrs Allen

PROJECT TITLE: ANALYSIS OF EXISTING DATA FROM THE SOCIAL SCIENCE COMPONENT OF A WHO PREGNANCY REGISTRY PILOT STUDY.

Thank you for submitting your study to the Faculty of Health Sciences Human Research Ethics Committee for review.

It s a pleasure to inform you that the HREC has formally approved the above-mentioned study.

Approval is granted for one year until the 30th January 2015.

Please submit a progress form, using the standardised Annual Report Form if the study continues beyond the approval period. Please submit a Standard Closure form if the study is completed within the approval period. (Forms can be found on our website: www.health.uct.ac.za/research/humanethics/forms)

Please note that the ongoing ethical conduct of the stucy remains the responsibility of the principal investigator.

Please quote the HREC reference no in all your correspondence.

Yours sincerely

PROFESSOR M BLOCKMAN
CHAIRPERSON, FHS HUMAN ETHICS
Federal Wide Assurance Number: FWA00001637.
Institutional Review Board (IRB) number: IRB00001938
This serves to confirm that the University of Cape Town Human Research Ethics Committee complies to the Ethics Standards for Clinical Research with a new drug in patients, based on the Medical Research Council (MRC-SA), Food and Drug Administration (FDA-USA), International Convention on Harmonisation Good Clinical Practice (ICH-GCP) and Declaration of Helsinki guidelines.
The Human Research Ethics Committee granting this approval is in compliance with the ICH Harmonised Tripartite Guidelines E6: Note for Guidance on Good Clinical Practice (CPMP/ICH/135/95) and FDA Code Federal Regulation Part 50, 56 and 312.

Appendix 2: Informed consent documents

Participant informed consent document in-depth interview (Chapter 5)

I would like to ask if you would consider talking to me about the study you are taking part in. We will sit down together for about an hour so that we can discuss your experiences over a drink and snack. I will ask you about how you came to be in the study and what happened during the study. I will also ask you about health and treatments in general and your own health and treatments in the past and since you started this study. This is to see that we do these studies in the best way possible. While related to the study you are already taking part in, this talk is part of a separate study, and, if you do not want to talk to me, it will not affect the first study in any way.

As this talk is quite long I need to tape record what we all say so that I don't forget anything. This recording will be written into a computer so that I can easily compare what you tell me with what other people tell me. I expect to talk to about 30 people in total.

You will only be identified by a code and not by name in any report that is published. People who might see this information includes others involved in this project, including the clinical trial team, representatives of the fund-holder, the ____ Research Ethics Committee (who has approved the plan for our talk and this form), the _____ and/or other foreign regulatory authorities and possibly your personal, or other, doctor if you have one. If you have questions about your rights in research, you may call _____ on _____.

It is entirely up to you whether you want to have this discussion or not and you can stop it any time without it affecting your care by the staff here or at any other clinic. We do not feel that you will have any risks in taking part in this talk. The benefits are that any new information you tell us could help some more with your medical care during this trial. We will cover transport costs you may have if you need to come back to the clinic or another place for this discussion with R150.

Thank you for taking the time to consider my request. If you would like, please take some time to think about it or discuss it with your relatives. Do you have any questions for me?

	Principal Investigator(s):	*Please initial each line*
1	I understand the information given to me by [ ] and in this form	____
2	I was given time to ask questions. I had these answered to my liking (delete if n/a.)	____
3	I was given time to consider this, and discuss taking part with others such as my family	____
4	I understand that I can stop at any time without my care or legal rights being affected	____
5	I understand what we discuss may be seen by other staff from this project, _____ Research Ethics Committee & regulatory authorities. My name will not appear in a report.	____
6	I agree freely to take part in the discussion described above	____

		dd	*mm*	*yyyy*
Signature: Participant	Print all initials and surname	Date		
		dd	*mm*	*yyyy*
Signature: Witness	Print all initials and surname	Date		
		dd	*mm*	*yyyy*
Signature: Interviewer	Print all initials and surname	Date		

Participant informed consent document focus group discussion (Chapter 5)

I would like to ask if you would consider talking to me and about 5 others about the study you are taking part in. We will sit together for about an hour so that we can discuss everyone's experiences and ideas over a drink and snack. I will ask the group about how you all heard about the study and common health problems and treatments [that people living with HIV have or take]. We will also discuss what experiences the group had during the study in terms of their own health and treatments. This is to see that we do these studies in the best way possible. While related to the study you are already taking part in, this talk is separate and, if you do not want to talk to me, it will not affect the other study in any way.

As this talk is quite long I need to tape record what we all say so that I don't forget anything. This recording will be written into a computer so that I can easily compare what you tell me with what other people tell me. I expect to talk to about 40 people in total.

You will only be identified by a number and not by name in any report and, in fact, anything you discuss will never be linked to the information you gave during the study unless you specifically would like to be referred back to the clinical trial team. It will also never form part of your medical records at this clinic or any clinic. People who might see the anonymous information includes others involved in this project and representatives of the fund-holder, the _____ Research Ethics Coordinating Committee (who has approved the plan for our talk and this form), local and/or other foreign regulatory authorities. If you have questions about your rights in research, you may call _____ on _____. For any question or advice on the implementation of this study, you may contact _____

It is up to you whether you want to have this discussion or not and you can stop it any time without it affecting your care by the staff here or at any other clinic. There is a risk that you may share some personal or confidential information by chance, with other members of the group. There is also a risk that you may feel uncomfortable talking about some of the topics. However, we do not wish for this to happen. **You do not have to answer any question or take part in the discussion if you feel the question(s) are too personal or if talking about them makes you uncomfortable.** There is no direct benefit to you for taking part but we hope it will help us conduct studies well for others in the future.

Thank you for taking the time to consider my request. If you would like, please take some time to think about it or discuss it with your relatives. We will cover transport costs you have as a result of coming to this discussion. Do you have any questions for me?

Principal Investigator(s): _____ *Please initial each line*

1 I understand the information given to me by [ ] and in this form _____

2 I was given time to ask questions. I had these answered to my liking (delete if n/a.) _____

3 I was given time to consider this, and discuss taking part with others such as my family _____

4 I understand that I can stop at any time without my care or legal rights being affected _____

5 I understand what we discuss may be seen by other staff from this project, [ ] Research Ethics Committee & regulatory authorities. My name will not appear in a report. _____

6 I agree freely to take part in the discussion described above _____

		dd *mm* *yyyy*
Signature: Participant	Print all initials and surname	Date
		dd *mm* *yyyy*
Signature: Witness	Print all initials and surname	Date
		dd *mm* *yyyy*
Signature: Interviewer	Print all initials and surname	Date

Staff informed consent document focus group discussion (Chapter 5)

I would like to ask if you would consider talking to me and other staff involved in the SEACAT/INTERACT study about your experiences with asking study participants about their medical/medication histories, adverse events and concomitants medications. We will sit together for about an hour so that we can discuss the group's experiences and ideas over a drink and snack. I will ask you to discuss with your colleagues what type and level (extent) of data you feel is necessary, potential barriers to obtaining accurate and complete patient reports of such data and, if necessary, then brainstorm ways to overcome these barriers. This is with a view to understanding whether we conduct our studies in regard to these issues in the best way possible.

As this talk is quite long I need to tape record what we all say so that I don't forget anything. This recording will be transcribed and entered into a specialist computer program for qualitative data so that I can organise the information and compare what is said with what others say. I expect to talk to about 12 people in total in the South African SEACAT studies and the Tanzanian INTERCAT studies.

You will only be identified by a number and not by your name in any report. People who might see this anonymous data includes others involved in this project, representatives of the fund-holder, and the _____ Research Ethics Committee (who has approved the protocol including this informed consent form). If you have questions about your rights in research, you may call _____ *on* _____.

It is up to you whether you want to have this discussion or not and you can stop it any time without it affecting your employment. There is a very small risk that you may share some personal or confidential information by chance, with other members of the group although this is not the intention of this research. You do not have to answer any question or take part in the discussion if you feel the question(s) are too personal or if talking about them makes you uncomfortable. The potential benefit to you for taking part is that this will help you to consider optimal ways for eliciting patient reported data but ultimately we hope the results will generally help those who are involved in this type of work.

Thank you for taking the time to consider my request. If you would like, please take some time to think about it or discuss it with others. Do you have any questions for me?

	Principal Investigator(s):	*Please initial each line*
1	I understand the information given to me by [ ] and in this form	____
2	I was given time to ask questions. I had these answered to my liking (delete if n/a.)	____
3	I was given time to consider this, and discuss taking part with others such as my family	____
4	I understand that I can stop at any time without my care or legal rights being affected	____
5	I understand what we discuss may be seen by other staff from this project, [UCT] Research Ethics Committee. My name will not appear in a report.	____
6	I agree freely to take part in the discussion described above	____

		dd	*mm*	*yyyy*
Signature: Participant	Print all initials and surname	Date		
		dd	*mm*	*yyyy*

Signature: Interviewer | Print all initials and surname | Date

Participant informed consent document focus group discussion (Chapter 6)

Appendix 5.3. INFORMATION AND CONSENT FORM FOR STUDY SUBJECTS
TAKING PART IN FOCUS GROUP DISCUSSIONS
(To be translated into the appropriate local languages)

Hello. We would like to ask you to consider taking part in a group discussion which will help us to improve the way we learn about the safety of medicines used by pregnant women.

During the discussion you and a few other women will meet with a person trained in this kind of work. This person will ask the group questions about how pregnant women manage their pregnancies. The group will also be asked about treatments usually taken for their and their unborn babies' health or to treat illness. We are also interested to hear about any medicines, herbs, injections, tablets, mixtures, powders or suppositories etc. that are recommended by health workers, friends or family members, traditional healers or birth attendants, or purchased from a shop or pharmacy. This discussion will take between one and a half to two hours of your time.

There is a risk that you may share some personal or confidential information by chance with other members of the group. There is also a risk that you may feel uncomfortable talking about some of the topics. However, we do not wish for this to happen. You do not have to answer any question or take part in the discussion if you feel the question(s) are too personal or if talking about them makes you uncomfortable.

The discussion will be recorded on a machine (tape recorder) and later written onto paper and put into a computer, however your name will not be identified in any report. We will keep these tapes for up to 5 years. These personal details will only be seen by staff arranging the discussion, other members of their team and groups who make sure the study is run properly. You are protected by strict guidelines enforced by the government in taking part in this discussion and two different groups called ethics committees of the (National Ethics Committee name) and the World Health Organization have also given their permission.

Your participation is completely voluntary. If you do not want to take part there will be no change to the service you receive from the health centre. You may stop taking part at any time. Taking part will not cost you or your family anything. We will provide for your transport costs from and back home at the approved rates to be paid and for you to receive a drink and snack during the discussion.

Do you have any questions for us?

I, (name of subject)..... confirm I understand and accept the 'information for study subjects' leaflet handed and explained to me in connection with this group discussion. I have had the opportunity to ask questions and I understand that the study records will be kept confidential. I am aware that I may withdraw my consent at any time without affecting my further health care.

Signature or Right Thumb Print: .....

Date: ..... Study ID: .....

Subject/Guardian

If Guardian, state relationship to recruited subject .....

Appendix 5-Informed consent forms V4

Name: .....

Signature or Right Thumb Print: .....

Date: .....

Witness (if applicable)

Name: .....

Signed: .....

Date .....

Informant (Study Staff Member)