Cochrane Library

Cochrane Database of Systematic Reviews

Eliciting adverse effects data from participants in clinical trials (Review)

Allen EN, Chandler CIR, Mandimika N, Leisegang C, Barnes K

Allen EN, Chandler CIR, Mandimika N, Leisegang C, Barnes K.
Eliciting adverse effects data from participants in clinical trials.
Cochrane Database of Systematic Reviews 2018, Issue 1. Art. No.: MR000039.
DOI: 10.1002/14651858.MR000039.pub2.

www.cochranelibrary.com

WILEY

TABLE OF CONTENTS

[Methodology Review]

Eliciting adverse effects data from participants in clinical trials

Elizabeth N Allen[1], Clare IR Chandler[2], Nyaradzo Mandimika[1], Cordelia Leisegang[1], Karen Barnes[1]

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

Contact address: Elizabeth N Allen, Division of Clinical Pharmacology, Department of Medicine, University of Cape Town, K45, Old Main Building, Groote Schuur Hospital, Observatory, Cape Town, Western Cape, 7925, South Africa. elizabeth.allen@uct.ac.za.

Editorial group: Cochrane Methodology Review Group.
Publication status and date: New, published in Issue 1, 2018.

Citation: Allen EN, Chandler CIR, Mandimika N, Leisegang C, Barnes K. Eliciting adverse effects data from participants in clinical trials. *Cochrane Database of Systematic Reviews* 2018, Issue 1. Art. No.: MR000039. DOI: 10.1002/14651858.MR000039.pub2.

ABSTRACT

Background

Analysis of drug safety in clinical trials involves assessing adverse events (AEs) individually 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 great reliance on reports from participants to detect subjective symptoms, where he/she is often the only source of information. There is no consensus on how these reports should be elicited, although it is known that questioning methods influence the extent and nature of data detected. This leaves room for measurement error and undermines comparisons between studies and pooled analyses. This review investigated comparisons of methods used in trials to elicit participant-reported AEs. This should contribute to knowledge about the methodological challenges and possible solutions for achieving better, or more consistent, AE ascertainment in trials.

Objectives

To systematically review the research that has compared methods used within clinical drug trials (or methods that would be specific for such trials) to elicit information about AEs defined in the protocol or in the planning for the trial.

Search methods

Databases (searched to March 2015 unless indicated otherwise) included: Embase; MEDLINE; MEDLINE in Process and Other Non-Indexed Citations; Cochrane Methodology Register (July 2012); Cochrane Central Register of Controlled Trials (February 2015); Cochrane Database of Systematic Reviews; Database of Abstracts of Reviews of Effects (January 2015); Health Technology Assessment database (January 2015); CINAHL; CAB Abstracts; BIOSIS (July 2013); Science Citation Index; Social Science Citation Index; Conference Proceedings Citation Index - Science. The search used thesaurus headings and synonyms for the following concepts: (A): Adverse events AND measurement; (B): Participants AND elicitation (also other synonyms for extraction of information about adverse effects from people); (C): Participants AND checklists (also other synonyms as for B). Pragmatic ways were used to limit the results whilst trying to maintain sensitivity. There were no date or sample size restrictions but only reports published in English were included fully, because of resource constraints as regards translation.

Selection criteria

Two types of studies were included: drug trials comparing two or more methods within- or between-participants to elicit participant-reported AEs, and research studies performed outside the context of a trial to compare methods which could be used in trials (evidenced by reference to such applicability). Primary outcome data included AEs elicited from participants taking part in any such clinical trial. We included any participant-reported data relevant for an assessment of drug-related harm, using the original 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 independently reviewed for final eligibility. Relevant data were extracted and subjected to a 100% check. Disagreements were resolved by discussion, involving a third author. The risk of bias was independently assessed by two authors. The Cochrane 'Risk of bias' tool was used for reports comparing outcomes between participants, while for within-participant comparisons, each study was critically evaluated in terms of potential impact of the design and conduct on findings using the framework of selection, performance, detection, attrition, reporting, and other biases. An attempt was made to contact authors to retrieve protocols or specific relevant missing information. Reports were not excluded on the basis of quality unless data for outcomes were impossible to compare (e.g. where denominators differed). A narrative synthesis was conducted because differences in study design and presentation meant that a quantitative meta-analysis was not possible.

Main results

The 33 eligible studies largely compared open questions with checklist-type questions or rating scales. Two included participant interviews. Despite different designs, populations and details of questioning methods, the narrative review showed that more specific questioning of participants led to more AEs detected compared to a more general enquiry. A subset of six studies suggested 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, two studies showed that quite severe or debilitating AEs were only detected by an interview, while other studies 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 groups. There was no common statistical rubric, but we were able to represent some effect measures as a risk ratio of the proportion of participants with at least one AE. This showed a lower level of reporting for open questions (O) compared to checklists (CL), with a range for the risk ratios of 0.12 to 0.64.

Authors' conclusions

This review supports concerns that methods to elicit participant-reported AEs influence the detection of these data. There was a risk for under-detection of AEs in studies using a more general elicitation method compared to those using a comprehensive method. These AEs may be important from a clinical perspective or for patients. This under-detection could compromise ability to pool AE data. However, the impact on the nature of the AE detected by different methods is unclear. The wide variety and low quality of methods to compare elicitation strategies limited this review. Future studies would be improved by using and reporting clear definitions and terminology for AEs (and other important variables), 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 endpoints in a 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 for use by both staff and participants.

PLAIN LANGUAGE SUMMARY

Questioning clinical trial participants about their health in order to collect information on adverse effects of drugs

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 doing other kinds of tests. The trial staff can also ask participants about how they are feeling after 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 some trials, participants may be asked a simple open question such as 'how have you been feeling?', while in other trials, participants may be asked about whether they have had any of a long list of possible symptoms (such as 'have you had a headache, stomach ache, or sore muscles?'). 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 might then affect the trial's results and what we know about the side effects of drugs.

We did this review to look at studies that compared different types of participant questioning methods in order to investigate these issues. We 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, drugs, and patients studied, we 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 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 were not reported when a participant is asked an open question and these severe symptoms will only be reported with the more specific question. This makes it difficult to say whether one method is better than any other and the different questioning methods may, in fact, be complementary and therefore should be used together. It is also difficult to say what a specific question should include, as it might take too long for a participant to have to answer a very long list. While more research is needed to resolve the remaining uncertainties, it is very important for trials to be clear about which kind of questioning was used when they publish their results. This will help readers understand the trial's findings about the side effects and make it easier to make accurate comparisons between trials.

BACKGROUND

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 risk profile in subsequent trials, particularly in under-studied population groups (ICH 2004). 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 (or sometimes adverse effects) (AEs) (ICH 1996). AEs are assessed either on an individual case basis or by aggregate statistical synthesis to provide evidence of likely adverse drug reactions (ADRs), which are those AEs that have a reasonable possibility of being caused by the trial drug (CIOMS 2005).

The processes involved in collecting, recording, analysing, and reporting AEs are generally considered more complex than those involved in evaluating the potential benefits of a drug, and methods are relatively less developed (Huang 2011). 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 (FDA 2005). For instance, studies have found that giving participants a checklist of potential AEs yields more reports than posing a general enquiry about change in health (Bent 2006). However, it is uncertain whether one way of questioning over another is better for detecting ADRs (Wernicke 2005). 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. This problem will also occur if trials use disparate methods. This restricts the ability to detect rare ADRs and to explore factors influencing the assessment of risk (FDA 2005; Huang 2011). This situation is compounded by generally poor reporting in subsequent publications about which methods were used to determine participant-reported AEs (Ioannidis 2004).

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. Studies that 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.

How these methods might work

Little is known specifically about how different methods of AE elicitation work, although this is likely to be similar to questions about other topics, in that response to questioning involves comprehension, judgement, recall from memory, and communication of the response (Tourangeau 1984). Our earlier qualitative study of barriers to accurate and complete reporting of harms data suggests that questioning detail and terminology influences participants' recognition of health issues and treatments. Moreover, we suggested that the perceived relative importance of health issues and treatments to the participant may be a factor (Allen 2011).

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 about 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.

OBJECTIVES

To systematically review the research that has compared the methods used within clinical drug trials (or that would be specific for such trials) to elicit information about the AEs that were defined in the protocol or in the planning for the trial.

METHODS

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 but which could be used in clinical trials (evidenced by reference to such applicability).

Types of data

AEs elicited from participants taking part in a clinical trial. For the purposes of this review, AEs are defined as those outcomes that were prespecified 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), recognizing that the trial itself might reveal that these are not actually increased in an intervention group compared with a control group. Concomitant medication and medical history data were also included in this review if the eligible study 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. 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. Thus, we included studies that reported participant-reported data relevant for an assessment of drug-related tolerability or harm, using the original 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 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.

Search methods for identification of studies

There was no date or sample size restrictions in the searches, but only reports published in English were searched for and 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 16^{th} 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 an unmanageablely large number of results, mainly 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 eligible studies for this review); and (ii) studies which mentioned in their abstract that they collected data from participants about adverse effects (i.e. thousands of studies that would not be eligible for this review).

The information specialist, with help from information colleagues, used pragmatic methods 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 Appendix 1 for the search strategies used in each database (presented with the results). We used Endnote X3 to collect, deduplicate, screen titles/abstracts, and record decisions on inclusion of papers.

Searching other resources

We supplemented the electronic searches by checking reference lists of included reports, some excluded reports that were relevant to the topic, and other reports known to the authors who are familiar with the research area (Horsley 2011), handsearching recent relevant topic-area conference abstracts (International Conference on Pharmacoepidemiology and Therapeutic Risk Management, International Society of Pharmacovigilance annual meeting) (Scherer 2007), and searching online libraries of theses/dissertations.

Data collection and analysis

Selection of studies

The first author (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 CL) independently reviewed the remaining titles and abstracts for eligibility according to the Criteria for considering studies for this review, as described in the protocol for this review (Allen 2013b). 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 were recorded along with the primary reason for non-inclusion.

Data extraction and management

One review author (EA) extracted data onto a data extraction form according to a prespecified list, with a second review author (CL) 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 with two reports and modified before being finalised as:

- Authors;
- Date published;
- Summary of study methods including any drug(s), indications/inclusion criteria, assessments(s), and duration of follow-up;
- Data (AE or equivalent with original 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 group, if relevant;
  - The number/proportion and/or nature of AEs as defined by the original authors of the trial, by study group, 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 us, as the reviewers.

Assessment of risk of bias in included studies

The risk of bias was independently assessed by two review authors (EA and KB), according to the Cochrane 'Risk of bias' tool, as far as was feasible in terms of the actual study design encountered (Cochrane 2011). 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 biases that we considered important.
It is acknowledged that a 'risk of bias' assessment is dependent on the completeness and quality of the original study report and we attempted to contact the original authors whose email addresses were available to retrieve protocols or specific relevant missing information (Young 2011). We did not exclude reports from this review on the basis of quality, unless insufficient data for the comparison were reported.

Measures of the effect of the methods

Effect measures from different methods were compared, where possible, by assessing an overlap in 95% confidence intervals (Golder 2011).

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

We sought to minimize the amount of missing data through contact with original authors, as mentioned above (Young 2011). 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, were reported.

Assessment of heterogeneity

As noted above, pooled estimates could not be calculated, so we did not follow our plan to assess heterogeneity using the Chi^2 test and I^2 statistic (Higgins 2002).

Assessment of reporting biases

As noted above, we could not calculate pooled estimates and so we could not follow our plan to assess reporting bias using a funnel plot (Sterne 2001).

Data synthesis

As a meta-analysis of included studies was not possible, given differences in study designs, interventions, and presentation, we conducted a narrative synthesis using recommendations by Popay 2006. One author (EA) first examined the included studies 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. We developed a narrative summary of the scope of the study designs in order to look at aspects of study design across studies, and the Characteristics of included studies table was 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, we did this using raw or summary data, where possible, in order to develop common quantitative rubrics. We examined the tabulation for relationships within and between the studies, with the aim of identifying variables that potentially moderated the effects. We followed the same process, where relevant, for items relating to the impact of different elicitation methods on the between-drug effects, the nature of AEs detected, and previous or concomitant medication and medical history data reported.

Subgroup analysis and investigation of heterogeneity

We did not conduct any meta-analyses and so no quantitative subgroup analyses or investigations of heterogeneity were done.

Sensitivity analysis

Likewise, because we did not conduct any meta-analyses, no sensitivity analyses were performed.

RESULTS

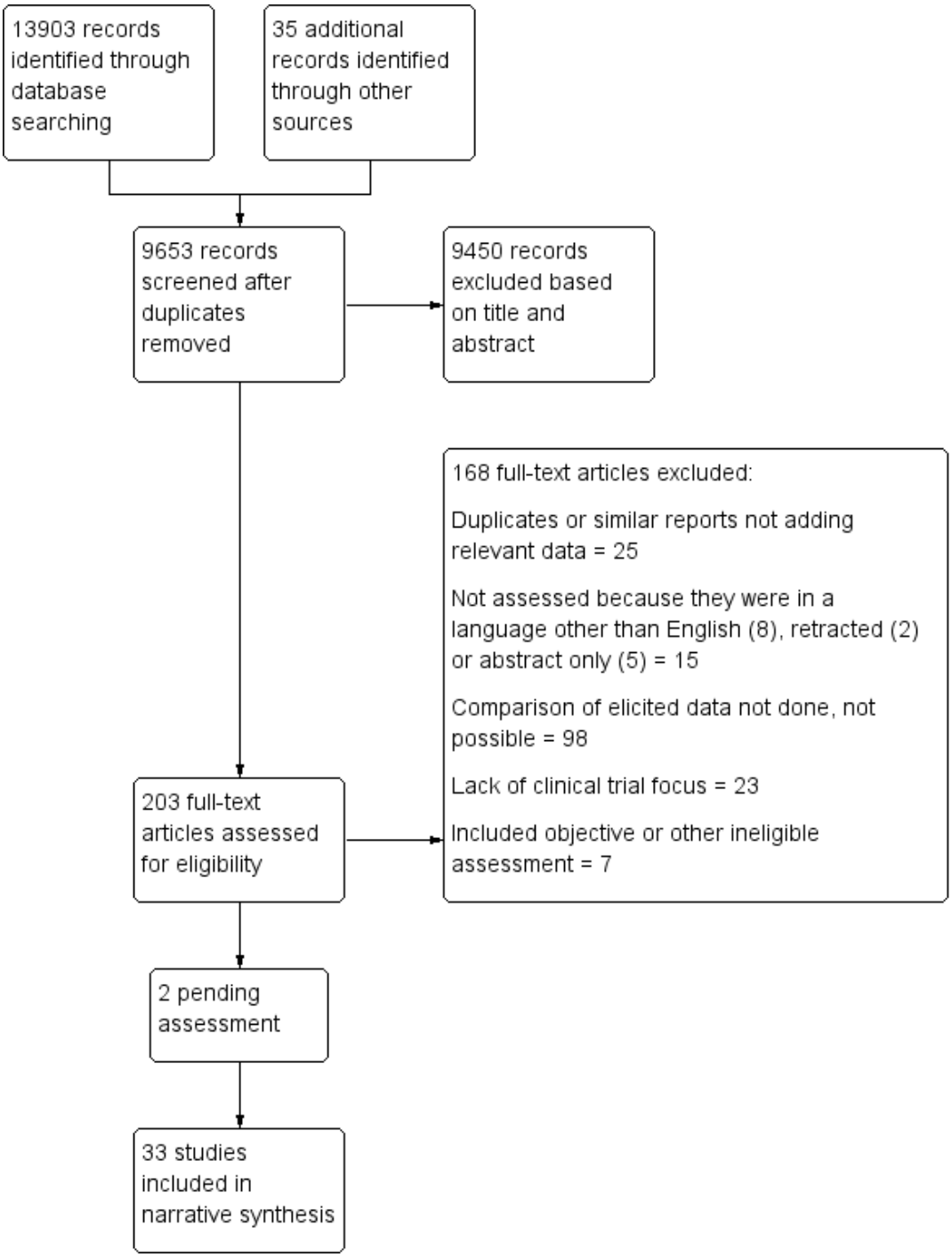
Description of studies

See the Characteristics of included studies, Characteristics of excluded studies and Characteristics of studies awaiting classification tables.

Results of the search

See Figure 1 for a flow diagram of the search metrics and Appendix 1 for the electronic search results. Our electronic search identified 13,903 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 handsearches. We sought the full text for 203 articles, of which 33 were eligible for inclusion. In total, 168 articles were excluded after checking the full text; 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 methods of 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. A further two articles are awaiting classification. See the Characteristics of excluded studies for details of articles which some readers may expect to be included. Nineteen included studies were found from the electronic search and 14 through non-electronic means.

Figure 1. Study flow diagram.



Included studies

Of the 33 eligible studies, 32 were published in full and one was a letter to the editor (Kruft 2007).

Theoretical basis for elicitation methods

The reports of the included studies were largely not explicit, or were unclear, about the theoretical basis for the work. Often, the expectation was simply that data were likely to be underreported if participants were not questioned in detail and that more specific questioning would be likely to increase the number of AE reports although it might have missed AEs that were not explicitly shown on the list or tool. One study, 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 (Rabkin 1992). Avery 1967 postulated that participants allowed to volunteer information may conceal or fail to recognise symptoms, while suggestible participants may report symptoms when questioned even though there is little objective evidence of the symptom being present. Rosenthal 1996 stated that 'connotation of words, among many other factors' could influence AE responses. Some authors 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). Where hypotheses were mentioned, they were largely based on a study or studies that we cited in this review, which may or may not be supported by the evidence. Studies that compared methods as part of a validation exercise for a new elicitation tool aimed to measure concordance (De Vries 2013; De Vries 2014). One study was explicit that it was conducted retrospectively with data already collected by different elicitation methods (Kruft 2007). Two were natural experiments, in that the method for eliciting AE data changed during a trial (Brent 2009; Monteiro 1987). For other studies, it was unclear whether the comparison was an a priori objective (Hermans 1994; Nicholls 1980; O'Connell 2007; Os 1994; Wernicke 2005; Yeo 1991).

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), psychiatry (Avery 1967, Brent 2009, Downing 1970, Greenhill 2004, Jacobson 1987, Landén 2005, Monteiro 1987, Rabkin 1992), ophthalmology (Barber 1995, Kruft 2007), diabetes (De Vries 2013, De Vries 2014), dysmenorrhoea (O'Connell 2007), gastrointestinal diseases (Barrowman 1970), gonorrhoea (Wallin 1981), malaria (Allen 2013) (a study conducted by the authors of this review), migraine (Sheftell 2004), Parkinson's disease (Perez-Lloret 2012), prostatic hyperplasia (Bent 2006), rheumatology (Huskisson 1974), and an unspecified indication for antihistamines (Lundberg 1980). Three studies were not related to any specific therapeutic area (Ciccolunghi 1975, Spilker 1987) 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. Two other studies were conducted in adolescents (Brent 2009, O'Connell 2007), one in children and adults (Wernicke 2005) and the remainder were conducted in adults (or we assumed them to be in adults when age was not explicit). However, because age ranges were missing from several papers, we did not include age in the Characteristics of included studies table.

Most studies were conducted in Europe (N = 17) or the United States (N = 12). Two other studies were multinational (Kruft 2007, Sheftell 2004), one African (Allen 2013) and the location of one was not reported (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 were nested within, or integral to, a trial. This latter group were all randomised trials, with either a reference drug or placebo, except for Allen 2013; and Wallin 1981, which used single-arm designs, and Török 1984, which used a mixture of randomised 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 studies of

specific AEs 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 our 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 our protocol (Allen 2013b) allowed for the inclusion of methods' studies conducted outside of a clinical trial, we expected that not all studies would collect or report AEs according to the International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use (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'). 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). The reports of four other studies were clear that they were collecting symptoms that were not necessarily treatment-emergent (Ciccolunghi 1975, Reilly 1992, Spilker 1987; and Török 1984). The remaining studies were unclear about this. For simplicity, we used AE as the default term hereafter. Rosenthal 1996 only reported on AEs that matched those on the checklist.

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 or symptom data, but there was not enough information to clarify whether measures of effect after baseline were treatment-emergent or not.

While we generally excluded studies that involved an objective assessment of AEs (e.g. through laboratory tests or physical examinations), 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 was possible that one of the methods involved a doctor 'filtering' participant reports (i.e. just reporting those he or 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 (N = 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 open questions (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), and how methods were developed and applied (e.g. verbal, written, or electronic) varied widely within these comparisons. See Table 1 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 studies compared data elicited on one visit occasion only, while other studies used data from multiple visit occasions, which 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.

Number of AEs reported

Several studies reported the number or proportion of AEs, or both, 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. However, three studies 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 and Barber 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 four 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 gastrointestinal 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 (by O) or by INT.

Statistical tests of effect by elicitation method, where used, mostly included Chi^2 and Mann-Whitney U tests. However, De Vries 2013 also used normal curve deviate statistics (Z value) for the measure of agreement between methods, and 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 changed over time and sensitivity to change, while Perez-Lloret 2012 analysed factors related to spontaneous reporting of AEs.

Seventeen studies presented comparative data by study group, 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 study groups. 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 the elicitation methods that they compared.

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 participant). Barber 1995 reported the average bother, level of activity limitation, satisfaction with medication, compliance, and global quality of life domain scores for those participants 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 1994 reported 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 participant, 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 clini-

cal 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 1992 investigated categories of AEs that were reported by participants through use of different questioning methods; Rabkin 1992 specifically looked at whether certain AEs were underreported through O (e.g. sexual dysfunction) by highlighting AEs that were reported more than five times as often by CL compared to O.

Other relevant outcome variables

Other outcomes reported for some studies 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 inter-rater 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 our search was broad out of necessity, we listed only those studies which a reader might reasonably expect to see among the included studies in the Characteristics of excluded studies table. 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 of eligibility (and no relevant paper was found), if there was no relevance to clinical trials, if there was no comparison of elicitation methods, or data were not 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 because the data collected through the different elicitation methods were categorised differently.

Risk of bias in included studies

See Table 2, Table 3, Table 4 and Table 5.

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 in this design. However, this potential for bias was unlikely to have impacted the results substantially because all participants were exposed to all methods. For two studies, the order that participants received the questions were 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 and therefore had a low risk of bias. Ciccolunghi 1975 used a predetermined allocation list and Spilker 1987 used a table of random numbers which also reduced the risk of bias. We judged the remaining five studies as either at high risk of bias because it was clear that 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 was a lower risk of bias as different staff were involved. However, the latter, in itself, raises the possibility of inconsistent recording of AEs. Ciccolunghi 1975 had a large proportion of non-responders to the invitation to take part in their study, which may have increased the risk of bias because 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 being compared in the study; 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 they help participants fill in the forms. It is not feasible to blind participants when they are reporting AEs, but for those taking part in studies where the comparison was between-participants, there may have been less risk of bias from knowing which questioning method was used, as it was less likely that participants were made aware of different methods being used in different groups/sites. For the within-participant comparisons, it was 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 were not independent comparisons. This was acknowledged by some study authors (Allen 2013, Greenhill 2004). The risk of bias for Wallander 1991 may have been the highest because 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 have prompted what to report in the other).

Incomplete outcome data

Most studies had a low risk of attrition bias. However, the risk of bias in this domain was unclear in nine studies and two were considered of higher risk as they had high dropout rates and we 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 method as participants took forms for both questioning methods home to complete and 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 staff decided whether or not they had completed the forms. 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 might be less likely to be completed than a shorter one).

Selective reporting

Twenty studies had a low risk of selective reporting and the risk of bias in this domain was unclear in 11. Two studies were deemed high risk as they presented only a selection of data; Rosenthal 1996 presented only spontaneously-elicited AEs if they matched questions on the questionnaire that they were being compared to, and Wernicke 2005 selected for comparison those AEs reflecting the same symptom in the spontaneous and solicited methods (in order to calculate ratios of the rate reported for 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, we were unclear whether the elicitation methods were applied in the same order for all participants, while for Os 1994, we were unclear as to how questioning methods were applied, making it impossible 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. Furthermore, 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 was unclear whether there may have been some other inherent biases. Finally, all studies were limited by a lack of a true gold standard against which to assess the data reported by participants.

Effect of methods

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

See Table 6 and Table 7 for details on the effects of the methods on the number of AEs reported for between-participant comparisons; and Table 8 and Table 9 for details on the effects of the methods on the number of AEs reported for within-participant comparisons.

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 (Table 6 and Table 7). There was no common statistical rubric, but we were able to represent some effect measures as a risk ratio of the proportion of participants with at least one AE. This showed a lower level of reporting for O compared to CL, with a range for the risk ratios of 0.12 to 0.64. 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 groups

Four studies presented data by study group (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 treatment group 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 predetermined 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 in a different context, for Ciccolunghi 1975, there was no difference in the detection of AEs between participants taking medication and those not taking medication when comparing the two types of O. There was a statistically significantly greater number of AEs

detected by CL compared to O in participants taking medications, but no such difference was 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 Table 8 for details on the elicitation of the number of AEs and Table 9 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 studies except 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 is likely to 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 fewer women than men reporting AEs by O but more women than men reporting AEs by R (the latter was 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, 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 groups

Common findings were identified when considering if the question method influenced the ability to detect differences between study groups. Nicholls 1980 and Rosenthal 1996 (drug-drug comparisons) and Downing 1970 (drug-placebo comparisons) showed a statistically significant difference between groups 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 groups. However, they also found that there were more statistically significant differences between trial groups by CL compared to O (nine versus five AEs). Landén 2005 (drug-drug) and O'Connell 2007 (drug-placebo) showed no difference between groups. The two studies that compared a CL with an R also had conflicting results: Lundberg 1980 found a difference between drugs by their CL but not by their R, while Os 1994 found no difference between these types of tools.

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

See Table 10 and Table 11 for details on the effects of the methods on the nature of AEs reported for between- and within-participant comparisons.

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 (Table 10). Brent 2009 found no difference between O and R for reporting of serious suicidal or nonsuicidal 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 elicited by CL were fatigue, headache, and nasal congestion, compared to headache, back or muscle pain, and nasal congestion by O, so there appeared 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), although Greenhill 2004 showed that their drug-specific review CL detected more moderate AEs compared to O (Table 11). 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, but 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 participants 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 noncompliance compared to those not reporting spontaneously; the average global quality of life scores increased as participants reported AEs spontaneously and discontinued therapy. The two studies that looked at the duration of AEs detected by questioning method (Hermans 1994, 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 concluded 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 suggested that the phrase used to introduce the questioning may have inadvertently suggested to participants that the latter were not the topic of the enquiry.

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 groups, not placebo. Greenhill 2004 also found that 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. The final tool was a questionnaire of 252 items and 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 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 participants who experienced an AE through their CL tool at either a MedDRA® organ class or specific level. The positive predictive value was especially low for a four-week recall period. Jacobson 1987 found good overall interrater reliability for detecting AEs using both their O and CL methods (best when raters were both present for the same participant consultation) but low interrater reliability for individual AEs and measures such as duration, severity, and functional impairment.

Qualitative

The one study (which we had conducted) that incorporated a qualitative analysis (Allen 2013) found that the CL and INT facilitated participants' recognition of health issues and treatments, and consideration of what to report. Information about AEs, medical history or non-study medicine use or both was sometimes not reported because participants forgot, it was considered irrelevant or insignificant, or they feared the consequences of reporting this. Some medicine names were not known, and answers to questions were sometimes considered inferior to the information that could be obtained from blood tests 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 site. 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.

DISCUSSION

Our Cochrane methodology 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: namely, how to accurately represent the often unknown adverse effects of a drug on a myriad subjective endpoints. The review itself was complicated by the diversity in the participant populations, designs, and elicitation methods used in the included studies. It is also difficult to ensure quality in this type of methodology research and in differentiating AE reports from disease-related symptoms. For instance, there may have been publication biases whereby studies that did not find any major differences between methods may have chosen not to report this. However, our review did 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 or bothersome AEs to an open enquiry. While the studies used disparate methods to assess these endpoints (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 or report it, or both. 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 were only 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 research.

The research that was nested within comparative drug trials also had mixed results when considering if the questioning method influenced the ability to detect differences in harm between trial groups. Nicholls 1980 and Rosenthal 1996 (drug-drug comparisons) and Downing 1970 (drug-placebo comparisons) showed a statistically significant difference between groups when using a checklist-type tool, and no such effect when using an open enquiry. In contrast, 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 groups. However, they also found that there were more statistically significant differences between trial groups by a checklist approach compared to an open enquiry (nine versus five AEs). Borghi 1984 and Landén 2005 (drug-drug) and O'Connell 2007 (drug-placebo) all showed no difference between questioning 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 host 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 study groups, or that the study groups being compared have similar safety profiles.

One response to our 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, especially preregistration trials, which aim to collect all AEs (i.e. high sensitivity with no provision for specificity relating to AE severity, bother from AEs, or 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 the 252 items developed by De Vries 2013. More research is therefore needed to explore the practicalities of using tools of different lengths and designs and to 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 would be to use different types of questioning depending on what is known about the safety profile of the drug, which may itself be a factor in the phase of development; with more comprehensive questioning early in the process and less comprehensive questioning as data builds about a favourable benefit:harm profile. Provision for this is made by the US Food and Drug Administration (FDA): manufacturers are allowed to make a case for limiting the safety data to serious AEs, for instance (FDA 2012). In addition, for the treatment of life-threatening illness, there may be a case for using less a sensitive enquiry and focusing 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 that are important to participants, 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 (Edgerly 2008). 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 effects may not have been identified as yet and this ADR will only be detected after many participants have been exposed to the drug (CIOMS 2005).

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 suggested 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, two studies showed that quite severe or debilitating AEs were only detected by an interview, while other studies did not find a difference in the nature of AEs between elicitation methods. No conclusions could be made regarding the impact of questioning method on the ability to detect a statistically significant difference between study groups, because the findings of the research we reviewed were inconsistent.

Overall completeness and applicability of evidence

Our review shows that, for a wide variety of populations, more AEs will be reported when participants are asked more comprehensive questions about their health. Some of the authors of the included studies hypothesised that such intensive questioning is suggestive, the implication being that participants will be made to report an AE that is not actually real. However, we 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 (De Vries 2014). More research 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 (Basch 2014). As a minimum, there is a need for more studies using interviews to understand reporting behaviour in a variety of contexts.

Another observation from our findings is that few tools other than checklists and scales were used, aside from one study that tested a diary. This could be an issue of our methodology, resulting in studies comparing such tools being missed. Alternatively, it could be that there are trials that use such tools but without any related 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 endpoints, including health-related quality of life measures, the methods for eliciting AEs are 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 questioning method used, which will contribute to consistency in trial conduct. As important is the need to provide sufficient detail of the elicitation method when reporting results, so as to 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 focused on AEs and we did not identify other studies that focused exclusively on those other variables.

Quality of the evidence

We did not use the Grading Recommendations, Assessment, Development and Evaluation (GRADE) approach in this review because it was not possible to conduct any meta-analysis. The quality of evidence was limited by the heterogeneity of studies included, the design limitations inherent for this kind of methodology research and 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 our being inclusive in anticipation that the number of eligible studies may be low, these differences did limit our 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 relevant 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 when elicita-

tion methods were used in sequence, since the data reports were 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 reflected the impact of using more than one method together, compared to one of the methods being used alone. All studies were limited by a lack of a true gold standard against which to assess the data reported by participants. Most studies had a low or unclear risk of attrition bias, but two could be considered of higher risk because it was not clear if dropouts were related to the questioning method that they used (Ciccolunghi 1975, Wallander 1991).

Potential biases in the review process

There were practical reasons for limiting the search to studies published in English and reporting terms synonymous with AE three or more times in the title or abstract. However, the review could have been improved by extending the search to languages other than English. Our findings also suggested that the electronic search missed several publications identified by non-electronic means. This highlights the importance of conducting a thorough review of reference lists of both included studies, and other articles that are relevant to the topic (as we did for this review). This raises the possibility that we may have missed other eligible studies, which may have weakened our overall conclusions. However, changes to the search strategy are unlikely to have overcome these issues without increasing the number of references to an unmanageable level and this issue could be explored in further research into the review methodology, such as those suggested within the Study Within A Review initiative (Anon 2012).

While we excluded studies with objective measures of AEs if the subjective data could not be extracted separately, some included studies may have also included objectively measured AEs without reporting them as such. We included populations taking part in clinical trials and methods studies outside of a clinical trial, 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 (Allen 2013, Heaven 2006, Paterson 2008, Scott 2011). This review focused on clinical trials and methods studies relevant for trials. However, as trials are not the only valuable source of information of harms, and in fact have inherent limitations on the detection of harms, they may not always be the best method to evaluate different elicitation methods. Other types of studies should be also ideally be explored in this regard. Separating the three-way comparisons of within-participant studies into two-way comparisons may have distorted results due to the possible effects of priming of one method on a subsequent one. In addition, the number of AEs or ADRs in a particular study or population may have impacted the ability to detect a difference between questioning methods and the types of items selected for a checklist may also influence this. We chose to include studies regardless of whether they prospectively or retrospectively addressed the comparison of AE elicitation methods (and this was not always clear from the reports of studies), although in general it is less optimal to retrospectively report outcomes. Lastly, we were not always able to confirm the calculations of the original authors.

Agreements and disagreements with other studies or reviews

No similar reviews were identified.

AUTHORS' CONCLUSIONS

Implication for methodological research

The wide variety and inherent low quality of methods used to compare elicitation strategies in clinical drug trials limited this review. Although we recognise that this is 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 is the case with similar work conducted in other areas of pharmacoepidemiology, this research is hampered by the lack of a true gold standard against which to assess the data reported by participants. This means that measures are likely to be concordant rather than valid (West 2005). 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 endpoints in any 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 (FDA 2009).

ACKNOWLEDGEMENTS

We should like to acknowledge the Cochrane Adverse Effects Methods Group for their peer review of the protocol and this full review. We also thank Clare Chandler (who subsequently joined the author team) and Mike Clarke for reviewing aspects of the protocol during development, and Tamara Kredo for her comments during the data analysis. In particular, we would like to sincerely thank Fiona Beyer for her design and conduct of the electronic searches.

REFERENCES

References to studies included in this review

Allen 2013 *{published data only}*
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 trials. *BMC Medical Research Methodology* 2013;**13**:140.

Avery 1967 *{published data only}*
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.

Barber 1995 *{published data only}*
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.

Barrowman 1970 *{published data only}*
Barrowman JA, Herxheimer A, Kits TP. Unwanted effects of pentagastrin. *Clinical Pharmacology and Therapeutics* 1970;**11**(6):862–8.

Bent 2006 *{published data only}*
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.

Borghi 1984 *{published data only}*
Borghi C, Pallavini G, Comi D, Grillo M, 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.

Brent 2009 *{published data only}*
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. *American Journal of Psychiatry* 2009;**166**(4):418–26.

Ciccolunghi 1975 *{published data only}*
Ciccolunghi SN, Chaudri HA. A methodological study of some factors influencing the reporting of symptoms. *Journal of Clinical Pharmacology* 1975;**15**(7):496–505.

De Vries 2013 *{published data only}*
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;**36**(9):765–77.

De Vries 2014 *{published data only}*
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* 2014;**23**(9):2439–45.

Downing 1970 *{published data only}*
Downing RW, Rickels K, Meyers F. Side reactions in neurotics. I. A comparison of two methods of assessment. *Journal of Clinical Pharmacology and the Journal of New Drugs* 1970;**10**(5):289–97.

Greenhill 2004 *{published data only}*
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.

Hermans 1994 *{published data only}*
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.

Huskisson 1974 *{published data only}*
Huskisson EC, Wojtulewski JA. Measurement of side effects of drugs. *British Medical Journal* 1974;**2**(5921):698–9.

Jacobson 1987 *{published data only}*
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.

Kruft 2007 *{published data only}*
Kruft B, Nelson LA, Stewart JA, Stewart WC. Adverse event reporting. *Ophthalmology* 2007;**114**(7):1420.

Landén 2005 *{published data only}*
Landén M, Högberg P, Thase ME. Incidence of sexual side effects in refractory depression during treatment with citalopram or paroxetine. *Journal of Clinical Psychiatry* 2005;**66**(1):100–6.

Lundberg 1980 *{published data only}*
Lundberg PK. Assessment of drugs' side effects: Visual Analogue Scale versus check-list format. *Perceptual and Motor Skills* 1980;**50**(3):1067–73.

Monteiro 1987 *{published data only}*
Monteiro WO, Noshirvani HF, Marks IM, Lelliott PT. Anorgasmia from clomipramine in obsessive-compulsive disorder. A controlled trial. *British Journal of Psychiatry* 1987;**151**:107–12.

Nicholls 1980 *{published data only}*
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.

O'Connell 2007 *{published data only}*
O'Connell K, Davis AR, Kerns J. Oral contraceptives: side effects and depression in adolescent girls. *Contraception* 2007;**75**(4):299–304.

Os 1994 *{published data only}*
Os I, Bratland B, Dahlöf B, Gisholt K, Syvertsen JO, Tretli S. Female preponderance for lisinopril-induced cough in hypertension. *American Journal of Hypertension* 1994;**7**(11): 1012–5.

Perez-Lloret 2012 *{published data only}*
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.

Rabkin 1992 *{published data only}*
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.

Reilly 1992 *{published data only}*
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.

Rosenthal 1996 *{published data only}*
Rosenthal J, Bahrmann H, Benkert K, Baumgart P, Bönner 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.

Sheftell 2004 *{published data only}*
Sheftell FD, Feleppa M, Tepper SJ, Rapoport AM, Ciannella L, Bigal M. Assessment of adverse events associated with triptans - methods of assessment influence the results. *Headache* 2004;**44**(10):978–82.

Spilker 1987 *{published data only}*
Spilker AV, Kessler JM. Comparison of symptoms elicited by checklist and fill-in-the-blank questionnaires. *Pharmacoepidemiology News* 1987;**22**:272–7.

Török 1984 *{published data only}*
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* 1984;**26**(5):555–62.

Wallander 1991 *{published data only}*
Wallander MA, Dimenäs E, Svärdsudd 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.

Wallin 1981 *{published data only}*
Wallin J, Sjövall J. Detection of adverse drug reactions in a clinical trial using two types of questioning. *Clinical Therapeutics* 1981;**3**(6):450–2.

Wernicke 2005 *{published data only}*
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.

Yeo 1991 *{published data only}*
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 Clinical Pharmacology* 1991;**31**(3):356–9.

References to studies excluded from this review

Anderson 1994 *{published data only}*
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.

Aspinall 2002 *{published data only}*
Aspinall MB, Whittle J, Aspinall SL, Maher RL, Good CG. Improving adverse-drug-reaction reporting in ambulatory care clinics at a Veterans Affairs hospital. *American Journal of Health-system Pharmacy* 2002;**59**(9):841–5.

Atherton 2012 *{published data only}*
Atherton PJ, Burger KN, Loprinzi CL, Neben Wittich MA, Miller RC, Jatoi A, et al. Using the Skindex-16 and common terminology criteria for adverse events to assess rash symptoms: results of a pooled-analysis (N0993). *Supportive Care in Cancer* 2012;**20**(8):1729–35.

Basch 2014 *{published data only}*
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 National Cancer Institute* 2014;**106**(9): 244. [DOI: doi.org/10.1093/jnci/dju244

Bennett 2012 *{published data only}*
Bennett BK, Park SB, Lin CSY, Friedlander ML, Kiernan MC, Goldstein D. Impact of oxaliplatin-induced neuropathy: a patient perspective. *Supportive Care in Cancer* 2012;**20**(11):2959–67.

Bergh 2013 *{published data only}*
Van den Bergh JP, Bouts ME, Van der Veer E, Van der Velde RY, Janssen MJ, Geusens PP, et al. 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.

Bonierbale 2003 *{published data only}*
Bonierbale M, Lançon C, Tignol J. The ELIXIR study: evaluation of sexual dysfunction in 4557 depressed patients in France. *Current Medical Research and Opinion* 2003;**19** (2):114–24.

Brown 2005 *{published data only}*
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. *Journal of Clinical Psychiatry* 2005;**66**(6):756–60.

Byerly 2006 *{published data only}*
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.

Carreno 2008 ***{published data only}***
Carreño M, Gil-Nagel A, Sánchez JC, Elices E, Serratosa JM, Salas-Puig J, et al. Strategies to detect adverse effects of antiepileptic drugs in clinical practice. *Epilepsy & Behavior* 2008;**13**(1):178–83.

Coolbrandt 2011 ***{published data only}***
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?. *European Journal of Oncology Nursing* 2011;**15**(2):130–6.

Costa 1979 ***{published data only}***
Costa FV, Ambrosioni E, Magnani B. Side effects of antihypertensive drugs. Incidence and methods of collection. *International Journal of Clinical Pharmacology and Biopharmacy* 1979;**17**(9):405–9.

De Smedt 2011 ***{published data only}***
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-sectional survey. *International Journal of Nursing Studies* 2011;**48**(12): 1540–50.

Downie 2006 ***{published data only}***
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.

Edwards 1996 ***{published data only}***
Edwards JG, Dinan TG, Waller DG, Greentree SG. Double-blind comparative study of the antidepressant, unwanted and cardiac effects of minaprine and amitriptyline. *British Journal of Clinical Pharmacology* 1996;**42**(4):491–8.

Emslie 2006 ***{published data only}***
Emslie G, Kratochvil C, Vitiello B, Silva S, Mayes T, McNulty S, et al. Treatment for adolescents with depression study (TADS): safety results. *Journal of the American Academy of Child and Adolescent Psychiatry* 2006;**45**(12): 1440–55.

Fisher 1990 ***{published data only}***
Fisher S, Bryant SG. Postmarketing surveillance: accuracy of patient drug attribution judgments. *Clinical Pharmacology and Therapeutics* 1990;**48**(1):102–7.

Gelenberg 2013 ***{published data only}***
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. *Journal of Nervous and Mental Disease* 2013;**201**(4):266–73.

Glaser 1954 ***{published data only}***
Glaser EM, Whittow GC. Experimental errors in clinical trials. *Clinical Science* 1954;**13**(2):199–210.

Greenblatt 1964 ***{published data only}***
Greenblatt M. Controls in clinical research. *Clinical Pharmacology and Therapeutics* 1964;**5**(6):864–70.

Hakobyan 2011 ***{published data only}***
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 and Quality of Life Outcomes* 2011;**9**:83.

Hanesse 1994 ***{published data only}***
Hanesse B, Legras B, Royer RJ, Guillemin F, Briancon S. Adverse drug reactions: comparison of two report methods. *Pharmacoepidemiology and Drug Safety* 1994;**3**:223–9.

Homsi 2006 ***{published data only}***
Homsi J, Walsh D, Rivera N, Rybicki LA, Nelson KA, Legrand SB, et al. Symptom evaluation in palliative medicine: patient report vs systematic assessment. *Supportive Care in Cancer* 2006;**14**(5):444–53.

Iverson 2011 ***{published data only}***
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 of degarelix or leuprolide. *Prostate Cancer and Prostatic Diseases* 2011;**14**(2):184–90.

Jarernsiripornkul 2009 ***{published data only}***
Jarernsiripornkul N, Kakaew W, Loalukkana W, Krska J. Adverse drug reaction monitoring: comparing doctor and patient reporting for new drugs. *Pharmacoepidemiology and Drug Safety* 2009;**18**(3):240–5.

Jonsson 2011 ***{published data only}***
Jonsson T, Christrup LL, Højsted J, Villesen HH, Albjerg TH, Ravn-Nielsen LV, et al. Symptoms and side effects in chronic non-cancer pain: patient report vs. systematic assessment. *Acta Anaesthesiologica Scandinavica* 2011;**55**(1): 69–74.

Lambert 2003 ***{published data only}***
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. *Human Psychopharmacology* 2003;**18**(5):405–11.

Love 1989 ***{published data only}***
Love RR, Leventhal H, Easterling DV, Nerenz DR. Side effects and emotional distress during cancer chemotherapy. *Cancer* 1989;**63**(3):604–12.

Makaranada 1995 ***{published data only}***
Makarananda K, Sriwatanakul K, Pothisiri P, Eamrungroj S, Charoonroj P, Pongprayoon U. Prospective study of adverse drug reactions of antihypertensive drugs in Thai outpatients. *Journal of the Medical Association of Thailand* 1985;**68**(5):229–36.

Martys 1982 *{published data only}*
Martys CR. Monitoring adverse reactions to antibiotics in general practice.. *Journal of Epidemiology and Community Health* 1982;**36**(3):224–7.

Mei 2006 *{published data only}*
Mei PA, Montenegro MA, Guerreiro MM, Guerreiro CA. Pharmacovigilance in epileptic patients using antiepileptic drugs. *Arquivos de Neuro-psiquiatria* 2006;**64**(2A):198–201.

Möller 2000 *{published data only}*
Möller HJ, Glaser K, Leverkus F, Göbel C. Double-blind, multicenter comparative study of sertraline versus amitriptyline in outpatients with major depression. *Pharmacopsychiatry* 2000;**33**(6):206–12.

Olsen 1999 *{published data only}*
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 Pressure* 1999;**8**(2):94–101.

Pandina 2007 *{published data only}*
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 and Adolescent Psychiatry and Mental Health* 2007;**1**(1):3. [DOI: doi.org/10.1186/1753-2000-1-3

Rynn 2015 *{published data only}*
Rynn MA, Walkup JT, Compton SN, Sakolsky DJ, Sherrill JT, Shen S, et al. Child/adolescent anxiety multimodal study: evaluating safety. *Journal of the American Academy of Child and Adolescent Psychiatry* 2015;**54**(3):180–90.

Sheikh 2013 *{published data only}*
Sheikh SI, Nestorov I, Russell H, O'Gorman J, Huang R, Milne GL, et al. Tolerability and pharmacokinetics of delayed-release dimethyl fumarate administered with and without aspirin in healthy volunteers. *Clinical Therapeutics* 2013;**35**(10):1582–94.

Thomsen 1997 *{published data only}*
Thomsen HS. Frequency of acute adverse events to a non-ionic low-osmolar contrast medium: the effect of verbal interview. *Pharmacology & Toxicology* 1997;**80**(2):108–10.

Tran 1997 *{published data only}*
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. *Journal of Clinical Psychiatry* 1997;**58**(5):205–11.

Trindade 1998 *{published data only}*
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.

Van Haecht 1990 *{published data only}*
Van Haecht CH, Vander Stichele R, Bogaert MG. Package inserts for antihypertensive drugs: use by the patients and impact on adverse drug reactions. *European Journal of Clinical Pharmacology* 1990;**39**(6):551–4.

Waddell 2008 *{published data only}*
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.

Yusufi 2007 *{published data only}*
Yusufi B, Mukherjee S, Flanagan R, Paton C, Dunn G, Page E, et al. Prevalence and nature of side effects during clozapine maintenance treatment and the relationship with clozapine dose and plasma concentration. *International Clinical Psychopharmacology* 2007;**22**(4):238–43.

References to studies awaiting assessment

AMIS 1980 *{published data only}*
Aspirin Myocardial Infarction Study research group. A randomized, controlled trial of aspirin in persons recovered from myocardial infarction. *JAMA* 1980;**243**:661–9.

Mothapo 2015 *{published data only}*
Mothapo KM, Schellekens A, Van Crevel R, Keuter M, Grintjes-Huisman K, Koopmans P, et al. Improvement of depression and anxiety after discontinuation of long-term Efavirenz treatment. *CNS & Neurological Disorders Drug Targets* 2015;**14**(6):811–8. PUBMED: 25808896]

Additional references

Allen 2011
Allen EN, Barnes KI, Mushi AM, 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**(Suppl 1):A10.

Allen 2013
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.

Anon 2012
Anon. Education section - Studies within a review (SWAR). *Journal of Evidence-based Medicine* 2012;**5**:188–9. [DOI: 10.1111/j.1756-5391.2012.01193.x

CIOMS 2005
Council for International Organizations of Medical Sciences. *Management of Safety Information from Clinical Trials. Report of CIOMS Working Group VI*. Geneva: CIOMS, 2005.

Cochrane 2011
Higgins JPT, Green S, editor(s). *Cochrane Handbook for Systematic Reviews of Interventions*. Vol. **5.1.0**, Chichester (UK): John Wiley & Sons, 2011.

Edgerly 2008
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.

FDA 2005
US Food, Drug Administration. Reviewer guidance: Conducting a clinical safety review of a new product application and preparing a report on the review. www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation /Guidances/ucm072974.pdf (accessed prior to 14 September 2017).

FDA 2009
US Food, Drug Administration. Guidance for Industry: patient-reported outcome measures: use in medical product development to support labeling claims.. www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM193282.pdf (accessed prior to 14 September 2017).

FDA 2012
US Food, Drug Administration. Draft guidance for Industry: determining the extent of safety data collection needed in late stage premarket and postapproval clinical investigations. www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM291158.pdf (accessed prior to 14 September 2017).

Golder 2011
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. [PUBMED: 21559325]

Heaven 2006
Heaven B, Murtach 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**: 260–70.

Higgins 2002
Higgins JP, Thompson SG. Quantifying heterogeneity in a meta-analysis. *Statistics in Medicine* 2002;**21**:1539–58.

Horsley 2011
Horsley T, Dingwall O, Sampson M. Checking reference lists to find additional studies for systematic reviews. *Cochrane Database of Systematic Reviews* 2011, Issue 8. [DOI: 10.1002/14651858.MR000026.pub2

Huang 2011
Huang HY, Andrews E, Jones J, Skoven ML, Tilson H. Pitfalls in meta-analyses on adverse events reported form clinical trials. *Pharmacoepidemiology and Drug Safety* 2011; **20**(10):1014–20.

ICH 1996
International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use (ICH). Guideline for Good Clinical Practice (Topic E6/R1). www.ich.org/products/guidelines/efficacy/article/efficacy-guidelines.html. (accessed prior to 14 September 2017).

ICH 2004
International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use (ICH). Pharmacovigilance planning: notice for guidance on planning pharmacovigilance activities (Topic E2E). www.ich.org/products/guidelines/efficacy/article/efficacy-guidelines.html. (accessed prior to 14 September 2017).

Ioannidis 2004
Ioannidis JP, Evans SJ, Gøtzsche PC, O'Neill RT, Altman DG, Schulz K, et al. CONSORT Group. Better reporting of harms in randomized trials: an extension of the CONSORT statement. *Annals of Internal Medicine* 2004; **141**(10):781–8.

Paterson 2008
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* 2008;**14**(2):199–208.

Popay 2006
Popay J, Roberts H, Sowden A, Petticrew M, Arai L, Rodgers M, et al. Guidance on the conduct of narrative synthesis in systematic reviews. Results of an ESRC funded research project (as supplied prior to 14 September 2017). Data on file 2006.

Scherer 2007
Scherer RW, Langenberg P, Von Elm E. Full publication of results initially presented in abstracts. *Cochrane Database of Systematic Reviews* 2007, Issue 2. [DOI: 10.1002/14651858.MR000005.pub3

Scott 2011
Scott C, Walker J, White P, Lewith G. Forging convictions: the effects of active participation in a clinical trial. *Social Science and Medicine* 2011;**72**(12):2041–8.

Sterne 2001
Sterne JAC, Egger M. Funnel plots for detecting bias in meta-analysis: guidelines on choice of axis. *Journal of Clinical Epidemiology* 2001;**54**(10):1046–55.

Tourangeau 1984
Tourangeau R. Cognitive sciences and survey methods. In: Jabine T, Straf M, Tanur J, Tourangeau R editor(s). *Cognitive aspects of survey methodology: building a bridge between disciplines*. Washington D.C.: National Academy Press, 1984.

West 2005
West SL, Strom BL, Poll C. Validity of pharmacoepidemiologic drug and diagnosis data. In: Strom BL editor(s). *Pharmacoepidemiology*. 4. Chichester (UK): John Wiley & Sons, 2005:709–65.

Young 2011
Young T, Hopewell S. Methods for obtaining unpublished data. *Cochrane Database of Systematic Reviews* 2011, Issue 11. [DOI: 10.1002/14651858.MR000027.pub2

References to other published versions of this review

Allen 2013b

Allen EN, Chandler CIR, Mandimika N, Barnes K. Eliciting adverse effects data from participants in clinical trials. *Cochrane Database of Systematic Reviews* 2013, Issue 5. [DOI: 10.1002/14651858.MR000039

* *Indicates the major publication for the study*

CHARACTERISTICS OF STUDIES

Characteristics of included studies *[ordered by study ID]*

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 to 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. Could 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

Avery 1967 (*Continued*)

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 captured 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 measured the extent to which these side effects and any associated limitations in routine living activities interfered with health-related quality of life, medication compliance, and participant 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.

Barrowman 1970 (*Continued*)

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 $\geq$ 1 AE by method ($\chi 2$). SAE description
Notes	Method 3: checklist developed after an 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 multicentre, 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 (suggested 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

Borghi 1984 (*Continued*)

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 behavior 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 behavior 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 versus 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

Ciccolunghi 1975 (*Continued*)

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	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 v 4.0, 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	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-ended 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 v 4.0, 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 or depression or both 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 symptomatology: "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, participant 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 $\geq$ 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

Greenhill 2004 (*Continued*)

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 checklist 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

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	Occular 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

Landén 2005 (*Continued*)

	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 organisation 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 Obsessive Compulsive Disorder and DSM-III enrolled in a double-blind RCT of clomipramine versus 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 Stuctured 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 versus 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 participant 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 oral contraceptives, 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)

Os 1994 (*Continued*)

	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 participant 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 under-reporting, 95% CI, binomial test for differences of AEs affecting > 10% of participants. 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-Parkinsons Disease drugs from a literature search critically reviewed by a group of PD and pharmacovigilance specialists for consensus: general, gastro-intestinal, 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 pretreatment 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

Rabkin 1992 (*Continued*)

	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 multicentre, randomised, double-blind, placebo-controlled trial of clentiazem (US, N = 92). Within-participant comparison of data elicited by 2 consecutive methods pretreatment 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 symptom severity index 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 $\geq$ 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

Rosenthal 1996 (*Continued*)

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 might reflect the participant's well being. 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 might reflect the participant's well-being. 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 of 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 chloranolol and for subjective gastro-intestinal related symptoms
Data	AEs.
Comparisons	ELICITATION METHOD 1 Complaints reported spontaneously by the participants (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 gastro-intestinal 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: "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, nonparametric permutation, multivariate tests with Pitman's in multivariate form - reporting before randomisation treated as confounding variable
Notes	Method 1: used in previous population studies, included depression, tension, head, heart, lung, metabolism, musculoskeletal system, GI and urinary tracts. Patients placed 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 (Swedwn, 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 was 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 to 9, AMDP-5: 47 items rated 0 to 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 pretreatment, 8 and 24 weeks/withdrawal)
Data	Side effect (cough).

Yeo 1991 (*Continued*)

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 ≥ 8 mm. Between-treatment differences using Kolmogorov-Smirnov 2-sample test, $\chi 2$ with Yates correction, Fisher's exact, Wilcoxen rand sum tests
Notes	

ACE:angiotensin−convertingenzyme

ADE:adversedrugevent

AE: adverse eventAMDP-5: Assessment and documentation of psychopathology
ANOVA: analysis of varience
ASPECT: Assessment of symptoms and psychological effects in cardiovascular therapy
BBAEQ-M: Barkley behavior and adverse events questionnaire-modified
BSR: body system review
CI: confidence interval
COMTOL: comparison on ophthalmic medications for tolerability
COSTART: Coding symbols for a thesaurus of adverse reaction terms
CTCAE: Common terminology ceriteria for adverse events
D: drug
DSI: drug-specific inquiry
DSM-III: Diagnostic and statistical manual of mental disorders 3rd edition
GI: general inquiry
MAOI: monoamine oxidase inhibitor
MedDRA: Medical dictionary for regulatory activities
NB: nota bene
no.: number
O: open-ended question
OR: odds ratio
P: placebo
QoL: quality of life
RCT: randomised controlled trial
S: structured
SAFTEE: Systematic assessment for treatment emergent events
SAE; serious adverse event
SD: standard deviation
SI: structured or systematic inquiry
SMURF: Safety monitoring uniform form
SSAP: subjective symptom assessment profile
SSRI: selective serotonin reuptake inhibitor
T1: time 1
T2: time 2
UKU: Udvalg for Kliniske Undersøgelser
VAS: visual analogue scale
$\chi 2$: chi-squared

Z: zeta (standard score)

Characteristics of excluded studies *[ordered by study ID]*

Study	Reason for exclusion
Anderson 1994	No comparison of elicited data
Aspinall 2002	Lack of clinical trial focus
Atherton 2012	Included objective measure(s)
Basch 2014	No comparison of elicited data
Bennett 2012	Assessed severity, experience of AEs already reported
Bergh 2013	Not possible to compare data between methods
Bonierbale 2003	Lack of clinical trial focus
Brown 2005	Included objective measure(s)
Byerly 2006	Lack of clinical trial focus
Carreno 2008	Lack of clinical trial focus
Coolbrandt 2011	Lack of clinical trial focus
Costa 1979	Lack of clinical trial focus
De Smedt 2011	Lack of clinical trial focus
Downie 2006	Lack of clinical trial focus
Edwards 1996	Incomplete relevant data reported
Emslie 2006	Not possible to compare data between methods
Fisher 1990	Lack of clinical trial focus
Gelenberg 2013	Not possible to compare data between methods
Glaser 1954	Not possible to compare data between methods
Greenblatt 1964	Incomplete relevant data reported
Hakobyan 2011	Lack of clinical trial focus

(*Continued*)

Hanesse 1994	Lack of clinical trial focus
Homsi 2006	Lack of clinical trial focus
Iverson 2011	Incomplete relevant data reported
Jarernsiripornkul 2009	Lack of clinical trial focus
Jonsson 2011	Lack of clinical trial focus
Lambert 2003	Included objective measure(s)
Love 1989	Lack of clinical trial focus
Makaranada 1995	Lack of clinical trial focus
Martys 1982	Lack of clinical trial focus
Mei 2006	Lack of clinical trial focus
Möller 2000	Not possible to compare data between methods
Olsen 1999	Lack of clinical trial focus
Pandina 2007	Included objective measure(s)
Rynn 2015	Not possible to compare data between methods
Sheikh 2013	Not possible to compare data between methods
Thomsen 1997	Lack of clinical trial focus
Tran 1997	Included objective measure(s)
Trindade 1998	Comparison between trials, not within trials
Van Haecht 1990	Lack of clinical trial focus
Waddell 2008	Lack of clinical trial focus
Yusufi 2007	Included objective measure(s)

Characteristics of studies awaiting assessment *[ordered by study ID]*

AMIS 1980

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 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..

AE:adverseevent

DASS: depression-anxiety-stress-scale
HIV: human immunodeficiency virus
OQ-45: Outcome Measure-45
SCL-90: Symptom Checklist-90

DATA AND ANALYSES

This review has no analyses.

ADDITIONAL TABLES

Table 1. Overview of questioning methods

	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	No details
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

Table 1. Overview of questioning methods (*Continued*)

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

Table 1. Overview of questioning methods (*Continued*)

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

Table 1. Overview of questioning methods (*Continued*)

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 participant 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

Table 1. Overview of questioning methods (*Continued*)

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 v 4.0, 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 v 4.0, 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 v 4.0, 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, participant 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

Table 1. Overview of questioning methods (*Continued*)

			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
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 organisation 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 oral contraceptives, 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	Predefined list of AEs	Verbal structured enquiry	Pre-defined list of most common ADRs to various anti-Parkinson's Disease drugs from a literature

Table 1. Overview of questioning methods (*Continued*)

			search critically reviewed by a group of PD and pharmacovigilance specialists 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. Participants placed 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 Symtpoms Scale (US National

Table 1. Overview of questioning methods (*Continued*)

			Institute of Mental Health) - 30 items including general symptoms such as trouble sleeping, diarrhoea, headache, trouble eating. BBAEQ-M: 24 items rated 0 to 9, AMDP-5: 47 items rated 0 to -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	Verbal interview	No details

Table 1. Overview of questioning methods (*Continued*)

	populations		
Monteiro 1987	Stuctured interview	Verbal interview	No details

ADE:adversedrugevent

ADR: adverse drug reaction
AE: adverse event
AMDP-5: Assessment and documentation of psychopathology
ASPECT: Assessment of symptoms and psychological effects in cardiovascular therapy
BBAEQ-M: Barkley behavior and adverse events questionnaire-modified
CTCAE: Common terminology ceriteria for adverse events
GI: general inquiry
MAOI: monoamine oxidase inhibitor
MedDRA: Medical dictionary for regulatory activities
SAFTEE: Systematic assessment for treatment emergent events
UKU: dvalg for Kliniske Undersøgelser
VAS: visual analogue scale

Table 2. Risk of bias (selection, performance and detection, and attrition) of included studies for within-participant comparisons

	Selection bias				**Performance and detection bias**		**Attrition bias**	
Author	**Random sequence generation**		**Allocation concealment**		**Blinding**		**Incomplete outcome data assessed**	
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	Appeared all participants completed all methods.
Barber 1995	N/A		High risk	Open by nature of study.	High risk	Participants/ assessors unlikely to be blinded due	Low risk	A few dropouts, unlikely related to methods.

Table 2. Risk of bias (selection, performance and detection, and attrition) of included studies for within-participant comparisons (*Continued*)

						to the nature of study. 'Possible 'priming" by earlier method		
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	Appeared all participants completed all methods.
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 dropouts, unlikely related to methods.
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.
Downing 1970	N/A		High risk	Open by nature of study.	High risk	Participants/ assessors unlikely to be blinded due to	Low risk	Appeared all participants completed all method.

Table 2. Risk of bias (selection, performance and detection, and attrition) of included studies for within-participant comparisons (*Continued*)

						the nature of study. 'Possible 'priming" by earlier method		
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	Appeared all participants completed all method.
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.
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.
Kruft 2007	N/A		High risk	Open by nature of study.	High risk	Participants/ assessors un-	Unclear	No information provided.

Table 2. Risk of bias (selection, performance and detection, and attrition) of included studies for within-participant comparisons (*Continued*)

						likely to be blinded due to the nature of study. Possible 'priming' by earlier method		
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	Appeared all participants completed all methods.
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	Appeared all participants completed all methods.
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 dropouts, unlikely related to methods.
Nicholls 1980	N/A		High risk	Open by nature of study.	High risk	Participants/ assessors unlikely to be	Unclear	Not clear why different numbers of par-

Table 2. Risk of bias (selection, performance and detection, and attrition) of included studies for within-participant comparisons (*Continued*)

						blinded due to the nature of study. Possible 'priming' by earlier method		ticipants for 2 methods
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	Appeared that all participants completed all methods.
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 dropouts but unlikely related to method.
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	Appeared that all participants completed all methods.
Rabkin 1992	N/A		High risk	Open by nature of study.	High risk	Participants/ assessors unlikely to be blinded due	Low risk	Appeared that all participants completed all methods.

Table 2. Risk of bias (selection, performance and detection, and attrition) of included studies for within-participant comparisons (*Continued*)

						to the nature of study. Possible 'priming' by earlier method		
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 dropouts but unlikely related to method.
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 fewer participants for method 2; could be related to matching of symptoms between methods
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	Appeared that all participants completed all methods.
Wallander 1991	N/A		High risk	Open by nature of study.	High risk	Participants/ assessors unlikely to be blinded due to	High risk	Significant number of dropouts (45/ 236), potentially related

Table 2. Risk of bias (selection, performance and detection, and attrition) of included studies for within-participant comparisons (*Continued*)

						the nature of study. Possible 'priming' by earlier method, especially as participants took forms for both questioning methods home to complete		to method (participants took forms for both questioning methods home to complete)
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.
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.
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

AE:adverseevent

N/A: not applicable
O: open question

Table 3. Risk of bias (reporting and other) of included studies for within-participant comparisons

	Reporting bias				Other biases	
Author	**No selective reporting**		**Explicit application**			
Allen 2013	Low risk	All data appeared presented.	Unclear	Possible variations in phraseology between participants.	N/A	
Barber 1995	Low risk	All data appeared presented.	Unclear	Possible variations in phraseology between participants. .	Unclear	Queried if methods applied in same order for all participants
Barrowman 1970	Low risk	All data appeared presented.	Low risk	Explicit instructions for staff to use.	N/A	
De Vries 2013	Low risk	All data appeared presented.	Low risk	Assumed explicit instructions for self-administration.	N/A	
De Vries 2014	Low risk	All data appeared presented.	Low risk	Assumed explicit instructions for self-administration.	N/A	
Downing 1970	Unclear	Only presented data that participants related to medication.	Low risk	Explicit instructions for staff to use.	N/A	
Greenhill 2004	Low risk	All data appeared presented.	Low risk	Explicit instructions for staff to use.	N/A	
Hermans 1994	Low risk	All data appeared presented.	Low risk	Assumed explicit instructions, particularly for self-administration	N/A	
Jacobson 1987	Unclear	Some overlap between data presented.	Low risk	Explicit instructions for staff to use.	N/A	

Table 3. Risk of bias (reporting and other) of included studies for within-participant comparisons (*Continued*)

Kruft 2007	Unclear	Difficult to ascertain as short report.	Unclear	Not clear from information provided.	Unclear	Meta-analysis
Landen 2005	Unclear	Summary data presented.	Low risk	Assumed explicit instructions, particularly for self-administration	N/A	
Lundberg 1980	Unclear	Summary data presented.	Low risk	Explicit instructions for staff to use.	N/A	
Montiero 1987	Unclear	Summary data presented.	Unclear	Not clear from information provided.	N/A	
Nicholls 1980	Low risk	All data appeared presented.	Unclear	Not clear from information provided.	N/A	
O'Connell 2007	Low risk	All data appeared presented.	Unclear	Not clear from information provided.	N/A	
Os 1994	Low risk	All data appeared presented.	Unclear	Not clear from information provided.	Unclear	Unclear order
Perez-Lloret 2012	Low risk	All data appeared presented.	Low risk	Explicit instructions for staff to use.	N/A	
Rabkin 1992	Low risk	All data appeared presented.	Low risk	Explicit instructions for staff to use.	N/A	
Reilly 1992	Low risk	All data appeared presented.	Unclear	Not clear from information provided.	N/A	
Rosenthal 1996	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	

Table 3. Risk of bias (reporting and other) of included studies for within-participant comparisons (*Continued*)

Sheftell 2004	Low risk	All data appeared presented.	Low risk	Assumed explicit instructions for self-administration.	N/A	
Wallander 1991	Low risk	All data appeared presented.	High risk	Although could 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	Unclear	Not clear from information provided.	Unclear	May have been variations in phraseology between participants in method 2	N/A	
Wernicke 2006	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	Low risk	All data appeared presented.	Low risk	Standard O and assumed explicit instructions for self-administration	N/A	

AE:adverseevent

N/A: not applicable
O: open question

Table 4. Risk of bias (selection, performance and detection, and attrition) of included studies for between-participant comparisons

	Selection bias				Performance and detection bias		Attrition bias	
Author	**Random sequence generation**		**Allocation concealment**		**Blinding**		**Incomplete outcome data assessed**	
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	Dropsouts were indicated but appeared similar between groups
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.
Borghi 1984	Unclear	No information provided.	Unclear	No information provided.	Low risk	Little information provided, how-	Low risk	Although group alloca-

Table 4. Risk of bias (selection, performance and detection, and attrition) of included studies for between-participant comparisons (*Continued*)

						ever the investigator was neither informed of the results of the self-reporting, nor did they help participants fill in the forms		tion of drop-outs unclear, there were only a few
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
Chiccolunghi 1975	Low risk	Predetermined 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 distributed in the internal mail and it was left to staff to decide whether to complete and return them, although similar numbers were returned for each elicitation type (57% versus 44%)
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	Unclear	No information provided.

Table 4. Risk of bias (selection, performance and detection, and attrition) of included studies for between-participant comparisons (*Continued*)

						blinded due to the nature of study; although groups were at different sites there may still have been room for biased assessments based on the method of questioning, and different staff may have elicited /recorded AEs differently		
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
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 have been room for biased assess-	Unclear	No information provided.

Table 4. Risk of bias (selection, performance and detection, and attrition) of included studies for between-participant comparisons (*Continued*)

						ments based on the method of questioning		

AE:adverseevent

CL: checklist
O: open question

Table 5. Risk of bias (reporting and other) of included studies for between-participant comparisons

	Reporting bias				**Other biases**	
Author	**No selective reporting**		**Explicit application**			
Avery 1967	Unclear	Summary data presented.	Low risk	Explicit phraseology for O. CL had specific items but not all was presented in laymans language so may have had inconsistent phraseology	N/A	
Bent 2006	Low risk	All data appeared presented.	Low risk	Assumed explicit instructions for self-administration.	N/A	
Borghi 1984	Low risk	All data appeared presented.	Unclear	Not clear from information provided.	N/A	
Brent 2009	Unclear	Brief summary data presented.	Unclear	Little information provided, although method 2 involved validated scales so assumed had explicit instructions on application	Unclear	Unclear whether method 1 was used in both groups.
Chiccolunghi 1975	Low risk	All data appeared presented.	Low risk	Assumed explicit instructions for self-administration.	Unclear	Large numbers of non-responders to invitation which could be related to methods

Table 5. Risk of bias (reporting and other) of included studies for between-participant comparisons (*Continued*)

Huskisson 1974	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	All data appeared presented.	Low risk	Assumed explicit instructions for self-administration.	Unclear	It was not stated how many were invited.
Torok 1984	Unclear	Relevant AE data only presented for participants taking chloranolol	Unclear	Not clear from information provided.	N/A	

AE:adverseevent

CL: checklist
O: open question

Table 6. Elicitation of the number of AEs reported for between-participant comparisons

Study	Therapy area	Endpoint	Follow-up	AEs elicited											
				Number participants				Number of AEs (total)				Number (%) participants with ≥1 AE			
				O	O	CL	R	O	O	CL	R	O	O	CL	R
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
Huskiss 1974**	Rheumtology	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
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
Bent 2006	Prostatic hy-	Any AE	< 1 - 24 months	70	70	74	N/A	11	14	238	N/A	10 (14)	9(13)	57 (77)	N/A

Table 6. Elicitation of the number of AEs reported for between-participant comparisons (*Continued*)

	per-plasia														
Borghi 1984*	Car-diol-ogy	Any AE	16 weeks	106	N/A	117	N/A	496	N/A	1556	N/A	Ox-prenolol 48 (45), chlortha done 60 (57)	N/A	Ox-prenolol 76 (65), chlortha done 81 (69)	N/A
Spilker 1987	No indi-ca-tion	Any AE	1 oc-ca-sion	132	N/A	166	N/A	229	N/A	581	N/A	106 (80)	N/A	154 (93)	N/A
Cic-col-unghi 1975	No indi-ca-tion	Any AE	1 oc-ca-sion	144	129	143	N/A	88	67	720	N/A	59 (41)	46 (36)	127 (89)	N/A
Brent 2009**	Psy-chia-try	Self-harm	12 weeks												
Suici-dal				181	N/A	N/A	153	N/A	N/A	N/A	N/A	16 (8.8)	N/A	N/A	32 (20.9)
Non-suici-dal				181	N/A	N/A	153	N/A	N/A	N/A	N/A	4 (2.2)	N/A	N/A	27 (17.6)
Sui-cide at-tempts				181	N/A	N/A	153	N/A	N/A	N/A	N/A	7 (3.9)	N/A	N/A	10 (6.5)

AE:adverseevent

CL:checklist

N/A:notapplicable

O:openquestion

vs:versus

R:ratingscale

**Allparticipantsasked AEsbyanopenquestion(O),butasthisprocesspossiblyinvolved′ filtering′ofreportsbythedoctor,thedatawerenotincludedinthereview*

***Huskissonusedacompositemeasureoff requencyandseverity*

****Aselectionofthetotal AEsarepresentedinthereviewasthoseobjectivelymeasuredwereexcluded*

†2−*sampletestofproportions*

††2−*samplettest*

Table 7. Effects of methods on the number of AEs reported for between-participant comparisons

Study	Therapy area	Endpoint	Follow-up	Effect (number of AEs)		
				Overall		By drug arm
				Proportion with ≥ 1 AE†	Description (including of other effect measure if different)	
Avery 1967	Psychiatry	Any AE	5 weeks	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	N/A	Unclear scoring; appears higher total score for AEs by CL than O (540 versus 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, gastrointestinal 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 gastrointestinal and 0.60 x more ('irrelevant') using O compared to CL. See paper for details
Torok 1984***	Cardiology	Any AE	Various	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	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

Table 7. Effects of methods on the number of AEs reported for between-participant comparisons (*Continued*)

Borghi 1984*	Cardiology	Any AE	16 weeks	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	-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	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				-0.12 (-0.22; 0.09) P = 0.0017	Group assigned to R reported significantly greater number of suicidal-related AEs than O. NB no completed suicides	N/A
Non-suicidal				-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

Table 7. Effects of methods on the number of AEs reported for between-participant comparisons (*Continued*)

Suicide attempts				-0.03 (-0.07; 0.02) P = 0.2689	No difference between R and O for reporting of suicide attempts	N/A

AE:adverseevent

CL:checklist

N/A:notapplicable

O:openquestion

vs:versus

R:ratingscale

∗AllparticipantsaskedAEsbyanopenquestion(O),butasthisprocesspossiblyinvolved′filtering′ofreportsbythedoctor,thedatawerenotincludedinthereview

∗∗Huskissonusedacompositemeasureoffrequencyandseverity

∗∗∗AselectionofthetotalAEsarepresentedinthereviewasthoseobjectivelymeasuredwereexcluded

†2−sampletestofproportions

††2−samplettest

Table 8. Elicitiation of the number of AEs reported for within-participant comparisons

Study	Therapy area	Endpoint	Follow-up	AEs elicited												
				Number of participants	Number of AEs (total)						Number (%) of participants with ≥1 AE					
					O	O	CL	CL	R	INT	O	O	CL	CL	R	INT
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
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
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

Table 8. Elicitiation of the number of AEs reported for within-participant comparisons (*Continued*)

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
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
Nichol 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
O'Con 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
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
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
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

Table 8. Elicitiation of the number of AEs reported for within-participant comparisons (*Continued*)

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
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
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
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
Sheftel 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
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
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
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
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
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
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

Table 8. Elicitiation of the number of AEs reported for within-participant comparisons (*Continued*)

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
Allen 2013*	Malari:	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
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
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

ADE:adversedrugevent

AE: adverse event
BSR: body system reviewCL: checklist
D: drug
DSI: drug-specific inquiry
INT: interview
N/A: not applicable
NK: not known
O: open question
O(B): blank page
PD:Parkinson'sDisease

PPV: positive predictive value
PSC: post-stroke control
R: rating scaleSD: standard deviation
SOC: system organ class
Sp-So: spontaneous-solicited index
VAS: visual analogue scale
W0, W2, W4, W8, W12:
χ^2: chi-squared
Z: zeta (standard score)

∗Twosites/groupsreportedseparatelyasdifferentparticipantpopulations

∗∗subsetofparticipantsfromCversusCLcomparisontheninterviewed

†McNemar'stestofproportions,††2−samplettest

Table 9. Effects of methods on the number of AEs reported for within-participant comparisons

Study	Therapy area	Endpoint	Follow-up	Effect (number of AEs)		
				Overall		By drug arm
				Test of proportions with ≥ 1 AE	Description (including of other effect measure if different)	
Barber 1995	Ophthalmology	Any AE	4 weeks	N/A	Average frequency of domain scores (number (%) mean (SD)) for those not reporting AEs by O but indicating AE by CL: ocular 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 ranged from participants experiencing AEs 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	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	OR 3.22 (1.49; 7.74) P = 0.0017†	While methods agreed for 85 (69%) of participants, CL elicited a significantly greater	More AEs reported with amplodipine versus isradipine when either O or CL used. No

Table 9. Effects of methods on the number of AEs reported for within-participant comparisons (*Continued*)

					number of AEs than O	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	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 to 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	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	N/A	See data by drug arm.	Statistically significant difference in mean number of AEs between drugs by CL

Table 9. Effects of methods on the number of AEs reported for within-participant comparisons (*Continued*)

						(5.4 labetolol, 3.6 propranolol P < 0. 05) but not by O (1. 6 labetolol, 1.5 propranolol)
O'Connell 2007	Dysmenorrhoea	Any AE	12 weeks	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	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 to 3.2)	N/A
Rabkin 1992	Psychiatry	Any AE	4 weeks	N/A	CL elicited 5 x mean number of AEs than O.	N/A
Reilly 1992	Cardiology	Any AE	10 weeks	W0: 0-.65 (-0.75; -0.54), P = 0.0000, W12: -0.47 (-0.59; -0.34), P = 0.0000 ††	% participants with ≥ 1 AE 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	-0.38 (-0.39; -0.36) P = 0.0000††	% of participants with ≥ 1AE was significantly greater with CL	Between-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	N/A	64% increase in number of	N/A

Table 9. Effects of methods on the number of AEs reported for within-participant comparisons (*Continued*)

					AEs elicited by CL after O.	
Wernicke 2006	Not known	Any AE	NK	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	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	-0.31 (-0.38; -0.25) P = 0.0000	Significantly more AEs reported through CL for those reporting 1, 2, 3 or more AEs	N/A
Landen 2005	Psychiatry	Sexual	4 weeks	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.	No statistically significant difference found between drugs by either method

Table 9. Effects of methods on the number of AEs reported for within-participant comparisons (*Continued*)

					01) and 29 women versus 20 men by R ($\chi^2$3.7, P = 0.06)	
Yeo 1991	Cardiology	Any AE	24 weeks	-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	Occular	NK	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	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	N/A	Number of AEs similar between the CLs (Z = -0.049, P = 0. 961)	N/A
Lundberg 1980	Antihistamine	Any AE	9 days	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

Table 9. Effects of methods on the number of AEs reported for within-participant comparisons (*Continued*)

Wallander 1991	Cardiology	Any AE	8 weeks	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	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	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	-0.17 (-0.20; -0.13) P = 0.0000	% participants with ≥ 1 AE significantly greater with CL compared to O. Not possible to compare frequency overall with R	Cough more frequent with lisinopril than nifedipine (8.5 versus 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 versus male with lisinopril (12.6 versus 4.4%, P = 0.0027), no difference for nifedipine. 3 fold difference with CL. By VAS, participant and spouse assessed frequency of lisino-

Table 9. Effects of methods on the number of AEs reported for within-participant comparisons (*Continued*)

						pril-associated AE similarly. With O, similar number of AEs independent of smoking, with CL a statistically significant difference between non-smokers and smokers (16 versus 7%, P = 0.0018)

ADE:adversedrugevent

AE: adverse event
BSR: body system reviewCL: checklist
D: drug
DSI: drug-specific inquiry
INT: interview
N/A: not applicable
NK: not known
O: open question
O(B): blank page
PD:Parkinson′sDisease

PPV: positive predictive value
PSC: post-stroke control
R: rating scaleSD: standard deviation
SOC: system organ class
Sp-So: spontaneous-solicited index
VAS: visual analogue scale
W0, W2, W4, W8, W12:
χ^2: chi-squared
Z: zeta (standard score)
∗Twosites/groupsreportedseparatelyasdifferentparticipantpopulations

∗∗subsetofparticipantsfromCversusCLcomparisonthenineterviewed

†McNemar′stestofproportions,††2−samplettest

Table 10. Effects of methods on the nature of AEs reported for between-participant comparisons

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

Table 10. Effects of methods on the nature of AEs reported for between-participant comparisons (*Continued*)

								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

Table 10. Effects of methods on the nature of AEs reported for between-participant comparisons (*Continued*)

								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% versus 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 versus 5 weeks (χ^2 = 9.41, df = 1, P = 0.004)
Non-suicidal								
Suicide attempts								

AE: adverse event

CL: checklist
df: degrees of freedom
N/A: not applicable
O: open question
R: rating scale
χ^2: chi-squared

**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 represented in review as those objectively measured were excluded*

Table 11. Effects of methods on the nature of AEs reported for within-participant comparisons

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

Table 11. Effects of methods on the nature of AEs reported for within-participant comparisons (*Continued*)

								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 limitations on QoL, more dissatisfaction with medication, and more noncompliance versus those not reporting by O
Barrowman 1970	Europe	Within	HV	Gastrointenstinal	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

Table 11. Effects of methods on the nature of AEs reported for within-participant comparisons (*Continued*)

								O and CL (26, 90%) had high mean intensity scores compared to those only reporting by CL (12, 50% - predominantly those taking tranquillizers) 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

Table 11. Effects of methods on the nature of AEs reported for within-participant comparisons (*Continued*)

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 with 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

Table 11. Effects of methods on the nature of AEs reported for within-participant comparisons (*Continued*)

								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 (P =

Table 11. Effects of methods on the nature of AEs reported for within-participant comparisons (*Continued*)

								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	Ocular	Unclear	NK	N/A
Montiero 1987	Europe	Within	Patients	Psychiatry	Sexual	Yes	12 weeks	N/A

Table 11. Effects of methods on the nature of AEs reported for within-participant comparisons (*Continued*)

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

AE:adverseevent

BSR: body system review
CL: checklist
DSI: drug-specific inquiry
INT: interview
N/A: not applicable
NB: nota bene
NK: not known
O: open question/spontaneous
QoL: quality of life
R: rating scale
*Two sites/groups reported separately as different participant populations

APPENDICES

Appendix 1. Appendix 1: final electronic search strategy and results

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

(*Continued*)

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 notification$)).ti	1952
14	((adverse or side or undesirable or treatment emergent or treatment related) adj2 (effect? or reaction? or event? or outcome? or symptom?) adj3 (measur$ 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

(*Continued*)

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

(*Continued*)

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

(*Continued*)

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	112

(*Continued*)

	notifie? or notification$)).ti	
11	((adverse or side or undesirable or treatment emergent or treatment related) adj2 (effect? or reaction? or event? or outcome? or symptom?) adj3 (measur$ 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	870
S20	S12 OR S13 OR S14 OR S15 OR S16 OR S17 OR S18 OR S19	58,723

(*Continued*)

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*))	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*))	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*))	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*))	747
S15	AB ("spontaneous report*" or "self report*" or "participant report*" or "patient report*" or "subject report*" or "self administer*")	28,140
S14	TI ("spontaneous report*" or "self report*" or "participant report*" or "patient report*" or "subject report*" or "self administer*")	4,309
S13	AB ((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*))	4,297
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*))	548
S11	S1 OR S2 OR S3 OR S4 OR S5 OR S6 OR S7 OR S8 OR S9 OR S10	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*))	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*))	7

(*Continued*)

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*))	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*))	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 self report* or record* or identif* or collect* or notify or notifie# or notification*))	718
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*))	364
S4	TI (adr or adrs)	40
S3	TI ((adverse or side or undesirable or treatment emergent or treatment related) N2 (effect# or reaction# or event# or outcome# or symptom#))	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#))	2,176
S1	(MH "Adverse Drug Event")	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)
16^{th} 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

(*Continued*)

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

(*Continued*)

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 (not being updated)
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

CONTRIBUTIONS OF AUTHORS

EA wrote the protocol with input from KB, CC, and NM. EA and NM or CL independently screened titles, abstracts or full texts for eligibility and KB had input where there was a need for discussion. EA extracted the data which was checked by CL. EA and CL independently undertook the risk of bias assessment. EA conducted the analysis and review write-up with input from KB and CC. All authors approved the final analyses and report of the review.

DECLARATIONS OF INTEREST

The review authors have no interests to declare.

SOURCES OF SUPPORT

Internal sources

- No sources of support supplied

External sources

- Bill and Melinda Gates Foundation, USA.

This review was 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

DIFFERENCES BETWEEN PROTOCOL AND REVIEW

The scope of the review was clarified as regards to 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.

NOTES

None.