BMJ open accessible medical research

Evaluating the impact and use of Transparent Reporting of Evaluations with Non-randomised Designs (TREND) reporting guidelines

Thomas Fuller,[1] Mark Pearson,[2] Jaime L Peters,[1,2] Rob Anderson[2]

To cite: Fuller T, Pearson M, Peters JL, *et al*. Evaluating the impact and use of Transparent Reporting of Evaluations with Non-randomised Designs (TREND) reporting guidelines. *BMJ Open* 2012;**2**:e002073. doi:10.1136/bmjopen-2012-002073

► Prepublication history and additional material for this paper are available online. To view these files please visit the journal online (http://dx.doi.org/10.1136/bmjopen-2012-002073).

Received 6 September 2012
Revised 21 November 2012
Accepted 22 November 2012

[1]Peninsula Collaborations for Leadership in Applied Health Research and Care (PenCLAHRC), University of Exeter, Exeter, UK
[2]Peninsula Technology Assessment Group (PenTAG), University of Exeter, Exeter, UK

Correspondence to
Dr Thomas Fuller;
t.fuller@exeter.ac.uk

ABSTRACT

Introduction: Accurate and full reporting of evaluation of interventions in health research is needed for evidence synthesis and informed decision-making. Evidence suggests that biases and incomplete reporting affect the assessment of study validity and the ability to include this data in secondary research. The Transparent Reporting of Evaluations with Non-randomised Designs (TREND) reporting guideline was developed to improve the transparency and accuracy of the reporting of behavioural and public health evaluations with non-randomised designs. Evaluations of reporting guidelines have shown that they can be effective in improving reporting completeness. Although TREND occupies a niche within reporting guidelines, and despite it being 8 years since publication, no study yet has assessed its impact on reporting completeness or investigated what factors affect its use by authors and journal editors. This protocol describes two studies that aim to redress this.

Methods and analysis: Study 1 will use an observational design to examine the uptake and use of TREND by authors, and by journals in their instructions to authors. A comparison of reporting completeness and study quality of papers that do and do not use TREND to inform reporting will be made. Study 2 will use a cross-sectional survey to investigate what factors inhibit or facilitate authors' and journal editors' use of TREND. Semistructured interviews will also be conducted with a subset of authors and editors to explore findings from study 1 and the surveys in greater depth.

Ethics and dissemination: These studies will generate evidence of how implementation and dissemination of the TREND guideline has affected reporting completeness in studies with experimental, non-randomised designs within behavioural and public health research. The project has received ethics approval from the Research Ethics Committee of the Peninsula College of Medicine and Dentistry, Universities of Exeter and Plymouth.

ARTICLE SUMMARY

Article focus

Two studies using mixed research method will be conducted to examine

- How Transparent Reporting of Evaluations with Non-randomised Designs (TREND) is used by authors of research articles and by journals in their 'instructions to authors'.
- What the impact of TREND has been on the reporting completeness of articles which have *used* TREND.
- What factors affect authors' and journal editors' use of TREND and other reporting guidelines in behavioural and public health evaluations.

Key messages

- Reporting guidelines can be used to redress incomplete reporting of health research, but need to be used by authors before they can have any impact.
- The proposed studies will, for the first time, provide a description of where, when and how TREND has been used; an analysis and evaluation of the impact that TREND has had on reporting completeness of studies with non-randomised designs; and, generate primary research evidence on factors that affect authors and editors' use of reporting guidelines.

Strength and limitations of this study

- Use of multiple databases to locate studies that have used TREND, but are dependent on authors to cite the original TREND paper to be included and not use the checklist to misleadingly report the conduct of their study.
- Mixed methods research to collect primary, quantitative and indepth qualitative data on factors that affect the use of reporting guidelines.
- Selecting studies only in English limits generalisability.

INTRODUCTION

It is frequently acknowledged that the quality and completeness of reporting of health and medical research is highly varied.[1] As there is no consensus on definition of 'quality'", nor on how to assess the quality of the design and reporting of a study, it is clear that studies that are not described in sufficient

detail cause many practical and ethical problems for researchers and policy makers.[2] For example, a poorly reported study makes it impossible for it to be replicated or for the results to be compared with the existing knowledge, generalised to other populations or its data included in systematic reviews and meta-analyses.

Reporting guidelines were conceived of as an innovation designed to redress incomplete and/or variable quality of reporting of research. A reporting guideline is a checklist, flow diagram or explicit text to guide authors in reporting a specific type of research, developed using explicit methodology.[3] Over 190 reporting guidelines have been developed since 1996 when the guidelines for randomised controlled trials (RCTs)—Consolidated Standards of Reporting Trials (CONSORT)—were first published.[4] Reporting guidelines have since been developed for many specific types of study designs such as non-randomised designs (Transparent Reporting of Evaluations with Non-randomised Designs (TREND)[5]), observational studies (Strengthening Reporting of Observational studies in Epidemiology (STROBE)[6]), specialisations such as diagnostic test accuracy (Standards for the Reporting of Diagnostic Test Accuracy studies (STARD)[7]) and interventions in acupuncture (Standards for Reporting Interventions in Clinical Acupuncture Trials (STRICTA)).[8] More recently, the Enhancing Quality and Transparency of Health Research Network (EQUATOR) was established in 2006 to further support the promotion, dissemination and development of reporting guidelines.

Use of reporting guidelines

Although views differ on who should use reporting guidelines, they can be and are used by a range of people in a variety of roles including: experienced and inexperienced authors, journals, publishers, journal editors and article reviewers (though not as a quality-assessment tool).[9] For example, if reporting guidelines are applied by an author in the course of writing up a study, the reporting guideline can: help remove ambiguity from what was actually done and hence facilitate or enable replication of the study, enable an accurate assessment of biases that might have affected the results and enable data extraction from the study for meta-analyses.[10] Similarly, how a reporting guideline is 'used' also depends on the role and purpose the person has for applying or referring to the guideline. A researcher might *use* a reporting guideline to help inform study design and reporting, whereas a journal editor might *use* the guideline as part of their instruction to authors and require a checklist to be submitted with the manuscript.[11] Use of reporting guidelines is thus person, role or context specific.

The following two brief quotes highlight fundamental issues relevant to the successful dissemination and uptake of reporting guidelines.

> Reporting guidelines are only a tool, and like all tools they help only if used.[10]

> Our (Implementation Science) instructions to authors already state our support for reporting guidelines, but it is worth restating here because many authors seem not to use them.[11]

O'Connor's quote makes the point that reporting guidelines are themselves inert and require deliberate application by, for example, an author or journal editor, before they can affect reporting completeness.[10] Eccles *et al*[11] also highlight that, though the journal *Implementation Science* has recommended the use of reporting guidelines by including a reference to them in the journal's instructions to authors; it is still up to the authors to apply the guidelines to their research—something that, apparently, most have not done. This begs the questions of why authors have not done so, and what would have increased the likelihood of them doing so? For the TREND guideline, these questions will be addressed in study 2.

Reporting guidelines and challenges of evaluating their impact

Despite the proliferation of reporting guidelines there have been relatively few studies investigating the impact or effect that they have had on reporting completeness and study quality.[12] Still, it is possible to assess or measure the impact of reporting guidelines in numerous ways. For example, the impact of reporting guidelines could be measured by examining their level of uptake (eg, how frequently they are 'used' by researchers or editors) or by assessing any differences in reporting completeness between papers that do and do not use reporting guidelines.

Evaluations of the impact of reporting guidelines on reporting quality need to consider three key issues. First, there is an absence of an agreed definition of reporting and study quality, and how a term is defined and operationalised has implications for how and what is measured in assessment of a study.[13] The 'quality' of a study cannot be measured by using the same criteria in all circumstances (hence, arguably, the proliferation of so many research-reporting guidelines). A decision on whether a study is of high or low 'quality' is typically reached by logic based on the accumulated presence of prespecified attributes. It is thus possible for two studies that are regarded as high 'quality' to differ in some of their specific attributes. Second, given that there is no universally agreed definition of 'quality', there is no 'gold standard' measure of 'quality'.[14] A range of methods have been used to assess reporting completeness and study quality, and the difficulties of doing so are well documented.[15] For example, Jüni *et al*[16] assessed study quality using 25 different measures to score 17 RCTs. Their study revealed that depending on the measure used, different ratings of quality were obtained for the same study. This result illustrated how different assessment tools might focus on different aspects of study design or measure it in different ways and thus

reach a different conclusion as to overall quality. Others in more recent papers have argued that it is actually more important to examine specific methodological aspects of a paper than to refer to a summary or global rating of quality.[17 18] Third, following the publication of reporting guidelines, it is reasonable to expect that there is a time lag between publication and uptake of the recommendations. Undertaking an evaluation prematurely in the course of the dissemination of the guideline is likely to yield results that show little or no effect on reporting completeness.[19] Following the first evaluation of CONSORT which covered the 2 years immediately following publication of the guidelines,[20] it has been recognised that allowing a period of at least 18–24 months following the publication of reporting guidelines is more appropriate.[18]

In consideration of these issues, the current studies will cover a period from 2000 to the present, and use two measures of study quality in addition to a measure of reporting completeness.

Example: the use and impact of the CONSORT reporting guidelines

The CONSORT reporting guidelines for RCTs have been revised several times, are frequently cited by authors, widely endorsed by health and medical journals, editorial associations and have been evaluated in numerous studies.[21] Studies that have examined the use of reporting guidelines have most frequently focused on if and how guidelines such as CONSORT are used by journals in their instructions to authors[22–24] and, less frequently, how journal editors and reviewers use guidelines.[25] Although CONSORT is widely supported (it is 'endorsed by over 50% of the core medical journals listed in the Abridged Index Medicus on PubMed'[22] within samples of high impact factor medical journals, it has been reported that less than half the journals included a specific mention or reference to CONSORT in their instructions to authors (in 2005 and 2008, respectively)[22 23] and more recently, that fewer adhere to the guidelines,[24] or advise peer reviewers to use the guidelines.[25]

Two early evaluations on the impact of the CONSORT guidelines on reporting completeness revealed inconsistent and smaller than anticipated effects in improvements in reporting of RCTs.[18 26] A more recent review however reported stronger evidence suggesting that CONSORT is associated with improved reporting of RCTs.[27]

Given that the CONSORT guidelines have been widely used and evaluated, they can serve as a benchmark for gauging the uptake and impact of other reporting guidelines.

The TREND reporting guideline

Although randomised controlled study designs are held to be the gold standard of comparative effectiveness research, the study design is not always appropriate or possible when evaluating behavioural and or public health interventions. Instead, evaluations using non-randomised designs are frequently used. Non-randomised designs are, however, more susceptible to biases affecting internal validity, and reports that do not fully and or transparently describe the content and context of a study (eg, details of an intervention) limit readers' ability to assess external validity. In 2004, Des Jarlais *et al*[5] published the TREND reporting guideline (hereafter referred to as 'TREND') and accompanying checklist to address the issue of incomplete reporting in studies of behavioural and public health interventions using non-randomised designs. The development of the guideline and checklist were based on the CONSORT (2001) reporting guidelines for RCTs. TREND focused on the reporting of "theories used and descriptions of intervention and comparison conditions, research design, and methods of adjusting for possible biases in evaluations that use nonrandomised designs"(ref. 5 p 361). TREND was to be used for

> ...intervention evaluation studies using nonrandomized designs, not for all research using nonrandomized designs. Intervention evaluation studies would necessarily include (1) a defined intervention that is being studied and (2) a research design that provides for an assessment of the efficacy or effectiveness of the intervention...(p 362)

A website maintained by Centre for Disease Control and Prevention (http://www.cdc.gov/trendstatement/) provides information about the guideline and serves as a portal through which researchers can provide feedback to the guideline developers. Despite this facility, published articles commenting on limitations of TREND, a statement on the website indicating that TREND will be revised periodically and that the CONSORT checklist (on which TREND was based) has been changed, no revisions or amendments to the TREND checklist have been published.[28 29]

The decision to examine the uptake and impact of TREND was based on a combination of three factors. First, although problems exist with reporting of non-randomised trials in public health,[30 31] there is an acknowledged need to include evidence from these evaluations into evidence synthesis.[2 29] TREND was developed to help increase the reporting completeness of such intervention studies and, in turn, make data extraction for evidence synthesis easier.[5] However, it is unknown whether the guideline has met this objective. (Study 1 will address the first part of this question.) Second, TREND occupies a niche within the repository of reporting guidelines in that they are specifically intended for experimental (rather than observational) studies, in which participants are allocated to an experimental condition using a non-random method such as allocation to an intervention depending on date of birth, date of admission to hospital or geographical location. The allocation of participants to an intervention condition by an investigator using a non-random method thus differentiates it from a study being observational—in which case STROBE guidelines could be applied—or

experimental and randomised—in which case, the CONSORT guidelines might be applied/followed. Last, it has been 8 years since the publication of TREND, and while other guidelines have been evaluated within a similar time period following publication, no study has examined the uptake and impact of TREND on reporting completeness. Without such a study, the utility of the guidelines remains unknown.

Aims and approach to evaluating the impact of the TREND reporting guideline

Results from two systematic reviews of the diffusion of innovations in health services will inform the approach to evaluating the impact of TREND (in this case, the 'innovation').[32 33] The results from these reviews have highlighted the need for implementation research to take into consideration a broad range of factors when evaluating an intervention. For example, evaluations should consider the importance of context[32] and investigate factors such as the formal and informal decision-making processes, and the impact of different professional groups' perspectives, knowledge base and skill set[33] on the successful diffusion of innovations.

In this project, developing a coherent picture of where and how TREND has been used will provide valuable contextual information, to the assessment of study design and reporting completeness (study 1). Subsequent questionnaires and interviews with authors and journal editors will, for example, yield information from different professionals/academic backgrounds on the decision-making process underlying decisions to use (or not) reporting guidelines (study 2).

The proposed project will address the following primary research questions in two studies:

1. How is TREND used by authors of research articles?
2. How is TREND used in journals' 'instructions to authors'?
3. What is the impact of TREND on the reporting completeness of articles which have used TREND?
4. What factors affect authors' and journal editors' use of TREND and other reporting guidelines in behavioural and public health evaluations?

Questions 1, 2 and 3 will be addressed in study 1, and question 4 in study 2. Overall, the proposed studies will add the following new information to the existing literature: a description of where, when and how TREND has been used; an analysis and evaluation of the impact that TREND has had on reporting completeness of non-randomised, experimental studies; and, it will gather primary research evidence on factors that affect authors' and editors' use of reporting guidelines.

STUDY 1: THE UPTAKE AND IMPACT ON REPORTING COMPLETENESS OF THE TREND REPORTING GUIDELINE

Aim

1. Describe the uptake or use of TREND by authors and its use by journals in the instructions to authors.
2. Assess the impact of TREND on reporting completeness and study quality in a sample of studies that do and do not use TREND.

It is hypothesised that articles that report using TREND will have higher levels of reporting completeness, but not necessarily study quality than a sample of papers that could have, but have not, indicated the use of TREND. In addition, it is hypothesised that differences in reporting completeness will be affected by the amount the paper used TREND and the impact factor of the journal in which it was published.

METHODS

Design

Study 1 will be conducted in two parts. The first part will be an exploratory study to examine the uptake of TREND and generate a description of where and how TREND is used by authors, and how TREND is used by a sample of journals by examining their instructions to authors. The second part of study 1 will use an observational method contrasting a sample of papers that do and do not use TREND to guide reporting (post-TREND publication).

Study inclusion and exclusion criteria

In the first instance, studies will be included if they are identified by Scopus or Web of Science as having cited Des Jarlais *et al*. A subset of papers will be identified and selected for the purpose of assessment of reporting completeness and quality if they *use* TREND and are published in English. For the purposes of this study, 'use' of TREND is defined and operationalised as the reference, within a published article, to TREND being referred to, to inform or guide the planning, conduct and/or reporting of the research study and results.

Limiting the sample to papers published in English will unfortunately limit the generalisability of the results.

Study identification

Part 1

The procedure for identifying papers that cite the original TREND publication will be as follows:

1. Studies that potentially use TREND will be identified by conducting a search, using the 'cited by' function for papers that cite the TREND reporting guideline by Des Jarlais *et al*. As it has been established previously that individual databases such as Scopus and Web of Science do not cover all citations,[34 35] both will be searched to increase the chances of identifying a more representative sample of papers that cite Des Jarlais *et al*.
2. An additional search for papers that use TREND but might not cite Des Jarlais *et al*, instead referring to the CDC website, will be conducted by using Issue Crawler. (Issue Crawler is a free-to-use computer program available at: https://www.issuecrawler.net/

that searches the Internet for references to specific URLs.)

3. The search results from each database will be combined. Duplicates will be identified and removed using the 'de-duplicate' function in EndnoteX5 and through a manual check of the combined search results. Where duplicate references are identified, the one with most bibliographic information will be retained.
4. Abstracts and titles for each of the papers that cite TREND will be screened for references to use of TREND, English language and the type of paper. The number of studies that cite Des Jarlais *et al* but are not in English (and thus excluded from this research) will be recorded and subsequently reported.

The first author, TF, will primarily be responsible for screening the articles that cite TREND for indications or descriptions of *use* of the guideline. Dual screening is not routinely planned due to the specific focus of the search, that is, only papers that cite TREND will be identified. If there is any uncertainty or ambiguity within a paper as to whether or not TREND has been used, MP and JP will be consulted. In the event that this does not resolve the uncertainty, the corresponding author of the paper will be contacted by email to clarify whether or not TREND has been used in any way during the conduct or reporting of the study.

5. If, in the abstract, there is no reference to TREND and the paper is, for example, an editorial, or letter (ie, it is not an evaluation or study with a non-randomised experimental design) the full-text article will not be retrieved. Data capturing the details of these papers will be recorded.
6. If there is or is not a reference to TREND in the abstract and the paper reports a study where TREND could possibly have been used the full-text article will be retrieved. A conservative approach will be adopted to decrease the likelihood of missing any studies that use TREND.
7. Following electronic retrieval of the full-text article, an electronic search within the article using the 'find' function will be conducted to establish how TREND has been used. The term 'TREND' will be used initially as the search term as this acronym is included in the original citation of the TREND guidelines. If the acronym is not found, a search for the point, where the Des Jarlais *et al* article was cited, will be conducted to ensure that any use of the guidelines was not made in a descriptive way.
8. The journals that have published studies that cited Des Jarlais *et al*, and journals listed as TREND supporters on the CDC website will be searched for reference to TREND in their instructions to authors.

Part 2

Papers that report using TREND to inform reporting will be retained from part 1 to form the 'intervention' sample.

A search for a comparison group of papers, published 2 years after TREND that would be able to, but do not, use TREND will be undertaken. (The delay of 2 years allows for changes in reporting that might naturally have occurred during the dissemination of TREND.) Searching for public health literature presents challenges owing to there being no specific public health databases and no standardised 'public health language'.[36] Similarly, there are no standardised requirements to refer to study design in headings or abstracts, and the term 'non-randomised' includes a range of study designs and is used relatively infrequently.[37] Despite these challenges, the search strategy for a sample of comparison studies will be guided by Furlan *et al*[38] who have demonstrated that it is possible to identify non-randomised studies using a focused search strategy.

The search terms referring to study design and field will be extracted from Des Jarlais *et al* 2004 and papers that use TREND. (The final search terms and strategy will be reported in a subsequent publication.) The search strategy will aim to strike a balance between the use of specific (eg, pre–post study) and generic or broad search terms (eg, health) which would probably be too exclusive or inclusive, respectively[39] resulting in search results that are too small or large.

The search will be conducted within MEDLINE, PsycINFO and EMBASE. Owing to limited resources, the search period will start from the period of 2006 and end on the present period, and publications in English.

Search results will be screened by title and abstract to exclude papers that do not fall within the scope of TREND. Twenty-five per cent of the search results will be randomly selected and double-screened by a second screener—this will be shared between MP and JP. Discrepancies between decisions regarding including or excluding the paper will be resolved through discussion. In the event that this does not resolve the discrepancy, a third opinion (from RA) will be sought regarding the relevance of the paper for inclusion in the pool of potential comparators.

An equal number of comparator articles (not citing TREND) will be randomly selected from the search as studies with equal group sizes are most efficient.[40] The total number of comparators thus will be limited by the total number of papers that use TREND (a subset of the papers that cite TREND) and will not be known until after screening of the papers that cite Des Jarlais *et al*. Consequently, it is possible that this study will be underpowered. We will undertake a post hoc power analysis to check this possibility and take the result into consideration when discussing the implications of the study.

All included papers will then be assessed for reporting completeness and study quality using the measures described below.

Outcomes

Primary outcome

The primary outcome measure will be reporting completeness of studies using experimental non-randomised

designs. Reporting completeness will be measured as a percentage of the total number of items from the TREND checklist reported in the study.

With regard to the assessing use of TREND by journals, this will be measured as a categorical variable indicating the presence or absence of a reference to TREND in the journal's instructions to authors.

Assessing reporting completeness

In order to minimise the chance of reaching a self-evident conclusion, that is, when using the TREND checklist to assess reporting completeness, one finds that studies that use TREND have higher levels of reporting completeness than studies that do not use TREND—we sought to locate a second measure of reporting completeness. Numerous checklists for assessing reporting completeness were considered in addition to TREND (eg, extensions to CONSORT; the STROBE checklist[6]; and, the Standards for Quality Improvement Reporting Excellence (SQUIRE).[41] However, though there is some overlap between checklists, they typically differed on the majority of items. For example, the differences between TREND and the SQUIRE and STROBE checklists related predominantly to focus of the field of study and the specific nature of their respective targeted study designs. We, therefore, decided to use only the TREND checklist to identify the presence or absence of information considered important for reporting of evaluations with non-randomised designs.

Although the TREND checklist is the only measure used to assess reporting completeness, we will also assess study quality (below). This will provide us with a second, albeit indirect, measure of reporting completeness. In using the study quality measures in addition to the TREND checklist, if information on a particular item cannot be determined from the study report, this can be interpreted and coded as representing a gap in reporting.

TREND checklist[5]*:* The TREND checklist is made up of a total of 21 'items' (58 including subitems) that refer to the title and abstract, introduction, methodology, results and discussion sections. Most, but not all, of the items on the checklist have more than one sub-item, and all require a response to indicate their presence ('yes') or absence ('no').

There are no instructions in the TREND statement on what to do if items in the checklist are not relevant to a study. For the purposes of this study, the project team agreed a priori that if items in the checklist are not relevant to a study, this would be noted and the denominator reduced accordingly when calculating the percentage of reporting completeness. Similarly, if there are multiple components to an item whereby it might be possible to indicate that a part is present (ie, 'yes') and another absent (ie 'no') an appropriate fraction of completeness will be given. For example, if there are two components to an item and only one is present, half a 'point' would be taken into consideration for calculation of reporting completeness.

It is important to note that the TREND checklist is not designed to generate an overall rating of quality of the study. However, one could argue that the more items from the checklist are present in the study, higher the reporting quality.

Secondary outcomes

Study quality as measured by the Effective Public Health Practice Project Quality Assessment Tool (EPHPP-QA)[42] and the Graphical appraisal tool for epidemiological studies (GATE)[43] revised by the National Institute for Health and Clinical Excellence (NICE) will be the main secondary outcome. It is recognised that if the use of TREND is related to improved reporting completeness, studies that do not use TREND are likely to have greater amounts of missing information which, in turn, would make study quality harder to assess. However, currently, there is no established method of adjusting the assessment of study quality allowing for reporting completeness. An approach could be developed or added to the project whereby corresponding authors are contacted and requested to provide the information not reported in the original publication. This approach would enable a more accurate assessment of study quality but is beyond the scope of this project.

Assessing study quality

The EPHPP-QA and revised GATE tools have been selected for their utility in assessing studies using a range of designs, and because there are differences in the range of questions and response options covered by each measure. The Newcastle-Ottawa Scale[44] was also considered for use in this study but deemed to be too narrow in its focus on case–control studies to be likely to be relevant to many studies included in the samples.

Revised GATE[45]: GATE was revised by NICE with a view to making it appropriate for assessing most of the study designs used in health-related interventions. The revised GATE has five sections covering: population, method of allocation to intervention and comparison, outcomes, analyses, summary and has a total of 28 items. Section 1 (population) is intended to assess external validity, and the remaining sections are intended to assess internal validity. Item response options include 'good', 'mixed', 'poor', 'not reported' and 'not applicable'. An overall rating of the study's internal and external validity is made by assessors.

EPHPP-QA[42]: The EPHPP-QA tool is designed to generate an overall rating of the quality of the research design. The first six of the eight components are used to inform a global rating on an ordinal scale. The EPHPP-QA tool has 21 items in 8 components assessing: selection bias, study design, confounders, blinding, data collection methods, withdrawals and dropouts, intervention integrity and analysis. The tool is intended to be used to generate an overall rating of the study—strong,

moderate or weak—which is derived from ratings of the six components of study quality. In addition, there are prompts on the scoring sheet for two reviewers to discuss their respective ratings of the paper before reaching an, agreed, overall rating of study quality.

The EPHPP-QA also has support and evidence for its construct and content validity, test–retest reliability[42] and is reportedly easy to use taking approximately 10–15 min to complete.[2]

If data relevant to an item on the revised GATE or EPHPP tools cannot be determined from the study report this will be recorded and used to inform the assessment of reporting completeness.

Data collection and analysis

Descriptive and multivariate (where relevant) s analyses will be undertaken using PASW Statistics 18. Details of the planned analyses for parts 1 and 2 of study 1 follow.

Part 1

Data will be collected on the following to map the characteristics of studies that use TREND:

- Number of citations of Des Jarlais *et al*[5] by year of publication.
- How authors report how they have applied the guideline to/within their work.
- Whether or not the authors have cited the original guideline reference, but not actually applied the guidelines to their study design, conduct or report.
- The field of the paper that TREND has been used in.
- How much of TREND has been applied.
- Study design.
- Country in which study was conducted.
- Description of size of the group that the intervention being evaluated (eg, one-to-one, group-based and community-wide).
- Characteristics of the target population (eg, age, socioeconomic status and ethnicity).
- Instruction by journals and/or journal editors to reviewers to use reporting guideline checklists to inform recommendations for an article's publication.
- Journal impact factor.

Descriptive statistics will be calculated to generate a portrait of the uptake and use of TREND.

Part 2

Prior to assessing reporting completeness and study quality, and in order to improve consistency between reviewers, the methods of applying the TREND checklist, EPHPP-QA and revised GATE will be discussed. One reviewer (TF) will assess all the included papers for reporting completeness and study quality. A randomly selected sample of 20–25% of all papers will be assessed by a second reviewer (MP or JP). A check of inter-rater reliability will be conducted and any significant discrepancies between reviewers will be resolved through discussion. If this does not resolve the discrepancy, a third reviewer will assess the paper and act as an adjudicator.

Details of publications such as journal title and year of publication will not be concealed from reviewers as there is no strong evidence to indicate that this affects assessment of reporting completeness and study quality.[18] We considered the possibility of blinding reviewers to whether or not TREND was used in a particular study as this could reduce the impact of biases during the assessment of reporting completeness and study quality phase. The project is, however, only funded for one effective full-time staff member (TF) and thus precludes the additional staff time required to oversee the process of blinding. TF will be responsible for undertaking the searches, screening and reviewing, whereas MP and JP will be involved in both screening and reviewing of papers. The multiple roles that the authors have thus make concealment of whether or not TREND was used impossible, and represents a limitation of the study.

Following the assessment of reporting completeness and study quality, data regarding a journal's current use of TREND in their 'instructions to authors' will be retrieved from the respective journal's website. As in the method described by Kunath *et al*,[4] if two or more journals refer to the same instructions to authors on the same publisher's website, they will be treated as independent journals.[46] Last, whether or not there is a requirement for completed checklists to be submitted, with journal papers, will be recorded.

The information collected will predominantly be quantitative, categorical data reflecting the presence or absence of information relating to the use of TREND and characteristics of the journals in which the papers appear. Where authors and journals' instructions to authors describe the use of TREND, the text will be copied verbatim and categorised.

Analysis of covariance and descriptive statistics will be used to examine the frequencies and differences that exist in reporting completeness between papers that do and do not use TREND. Details regarding journal impact factor, whether the journal refers to TREND in its instructions to authors, or is listed as a 'supporter' of TREND on the CDC website, and the number of sections or components of TREND that a paper uses will be recorded and used as covariates in the analyses.

In addition, descriptive statistics will be used to examine similarities and differences in the assessment of study quality with EPHPP-QA and revised GATE.

STUDY 2: FACTORS THAT AFFECT THE USE OF THE TREND REPORTING GUIDELINE

Aim

This second, exploratory, study aims to collect primary research data from a sample of authors and journal editors regarding factors that influence the use of TREND. We predict that a combination of personal (eg, perceptions of the benefits of using reporting guidelines), practical (eg, journal requirements) and 'environmental' (eg, supporting material regarding TREND)

factors will determine the use or otherwise of TREND. Furthermore, we expect that when (or if) journals require authors to indicate that they have adhered to reporting guidelines when submitting a manuscript for publication, that this will override any personal preference or beliefs regarding the merits of reporting guidelines.

METHODS

Design

A cross-sectional survey will investigate what factors inhibit or facilitate authors and journal editors' use of TREND. Semi-structured interviews will be conducted after the surveys to explore issues not suited to an electronic survey with limited range of response options. The methods for the survey and semi-structured interviews will be described, consecutively.

Author and editor survey

Sample and recruitment

A purposive sample of authors and journal editors will be recruited into study 2. The sample of the authors will comprise of the nominated corresponding author of papers that have used TREND and those that have not indicated use of TREND. Editors who are invited to participate in the study will be those who edit journals that have published the papers included in study 1. Editors and authors must also be fluent English speakers, and conduct research or edit journals in the fields of behavioural and public health (ie, they are within the targeted scope of TREND). The time period covered in study 2 will range from 2000 to the present.

Publicly available contact details of the corresponding authors and editors will be used to invite potential participants into the study. The initial email invitation will provide information about why they are being invited to participate, the degree of involvement required, an informed consent form and contact details of the lead investigator should they have any questions about participating in the study. Two weeks following the initial email invitation, the participants will be sent a second email with a copy of the questionnaire, an informed consent form attached and instructions on how to complete and return the survey. The maximum sample size will be limited by the number of studies that report using TREND and the comparators included in study 1.

A question within the survey will ask participants if they would like to take part in an additional, subsequent semi-structured interview.

No financial incentive to participate in the study will be offered. In order to maximise the response rate, participants will be emailed up to two reminders to complete and return the questionnaire.

Data collection procedures

The questionnaire is expected to take approximately 15 min to complete and will predominantly be quantitative in nature. Participants will use Likert-type scales with dropdown menus to indicate their responses. Open questions designed to capture qualitative data will also be included. Participants will be able to respond using free text fields to address these questions.

Measures

Self-report questionnaires will be developed for authors and journal editors, respectively. The questions will relate to perceived barriers and facilitators of the use of reporting guidelines and will be developed from the literature on reporting guidelines and diffusion of innovation, and the findings of study 1. Experienced staff within the Institute of Health Services Research (IHSR), University of Exeter Medical School, Exeter (UK), will review the structure and wording of the items to enhance content and face validity.

Questionnaires will be written in, and saved as Microsoft Excel files. The questionnaires will be pretested to evaluate whether potential respondents are likely to interpret questions in a consistent manner. Depending on the feedback received at this stage items might be reworded or removed. Pilot testing with a small convenience sample of authors and journal editors working within the IHSR and Peninsula Technology Assessment Group (PenTAG), University of Exeter Medical School, Exeter (UK), will take place. Pilot testing will further assess the relevance of items, administrative ease of completing the questionnaire and the time it takes to complete the questionnaire. Feedback at this stage of the development of the questionnaire will be taken into consideration and items further revised or removed as needed.

The exact content of the author and editor surveys will not be finalised until pilot testing and study 1 are complete. However, it is likely that the surveys will focus on five domains. Examples of questions for participants are provided below.

1. Demographics

Demographic information will be collected to provide descriptive information about the participants. Information to be collected will include details of: profession/discipline, field of study, if an author or field of expertise if an editor, experience/years of research experience/number of publications, age range, gender, country of work and membership of professional associations.

2. Knowledge of TREND

Participants could be asked: if they know of TREND or other reporting guidelines, and if so, when and how they learned of the reporting guidelines.

3. Use of the guidelines

Questions relating to the use of reporting guidelines could include the following:

- Whether they have used TREND or another reporting guideline, and if so, how (eg, to help inform the reporting of their study);

- If they have used TREND in an article or instructions to authors, who decided to apply/include the guideline;
- Who they think reporting guidelines should be used by and
- Whether they have provided feedback to the TREND group regarding their experiences of using the guideline.

4. Perceptions of the reporting guidelines and their impact

Participants could be asked:

- What they perceive to be the strengths and limitations of the reporting guidelines;
- How credible and transparent they think the guideline development process was;
- How useful and important reporting guidelines are for the reporting/instructions for authors of their own and others research/journals and
- How important for improving the reporting completeness of research they think it is to follow reporting guidelines such as TREND.

5. Factors that affect the use of TREND

Questions relating to factors that might affect participants' use of TREND and other reporting guidelines are likely to focus on what they perceived as barriers to the use of TREND (eg, lack of awareness of the guidelines, word limits, perceived weakness or gaps in the guidelines, no perceived need for them, inconsistent information regarding the guidelines and the requirement to use them), and perceived facilitators to the use of TREND (eg, journal requirements for article submission, knowledge and previous use of reporting guidelines, beliefs about the importance of following reporting guidelines and high-level editorial and author support for the use of guidelines).

Data management and analysis

Data from the questionnaires will be de-identified and entered into PASW 18 by TF. Prior to data analysis, data cleaning will occur and a random selection of 5% of surveys will be entered a second time to check for accuracy of data entry. Multiple imputation method within PASW 18 will be used to address missing data.

Descriptive statistics will be used to summarise participants' responses from the surveys. Reporting of the survey-based research will adhere to recommendations made by Kelly *et al*.[47]

Author and editor semi-structured interviews

Sample and recruitment

Approximately, 10 authors and editors who indicated a willingness to take part in a semi-structured telephone interview will be recruited. The authors and editors will be identified from their positive response to the survey question about their willingness to be interviewed. The authors and editors will be contacted via email to arrange a mutually convenient time to complete the semi-structured interview.

The transcription and thematic analysis of interviews will commence as soon as possible after the first interview to enable the project team to determine when and/or whether a thematic 'saturation point' is reached. In this way, the sample size can be increased if needed. A second call for volunteers to participate in an interview will be made if required.

Data collection procedures

Participants willing to be interviewed will be contacted via email to arrange a time to complete a semi-structured interview by telephone, Skype, or if in the same geographical location, face-to-face. Details of the context of the interview will be recorded. Using semi-structured interviews will enable the participants to elaborate on questions and responses obtained from the survey that had previously been completed. In addition, semi-structured interviews, with some standardised questions, increase the comparability of participants' responses.

The interviews will be digitally recorded and transcribed verbatim. Nvivo 9 software will be used to manage the data files and facilitate thematic analysis. Following transcription of the interviews, the participants will be offered the chance to review the transcripts to ensure that their views have been accurately expressed and transcribed. At this point, they could correct any errors in transcribing; add information not given at the time of the interview; and, remove sections they do not wish to be reported. Participants choosing to review the transcripts will be asked to do so within a 10-day period so as not to cause delays to the analysis and completion of the project.

Measures

The purpose of the semi-structured interview will be to obtain more detailed and background information for making decisions regarding the use of reporting guidelines. The results of study 1, the survey in study 2 and the work of Greenhalgh *et al*[32] will inform the topics of questions in the interview. It is anticipated, for example, that the questions will address the context and factors such as informal decision-making processes, and the impact of different professional groups' perspectives, knowledge base and skill set on the use of reporting guidelines.

Data management and analysis

Digital recordings of interviews will be stored securely on the University of Exeter network and thepassword protected. Data will be de-identified and any hardcopies of transcriptions will either be stored in locked filing cabinets within locked rooms or disposed of using confidential document disposal.

Thematic analysis will be undertaken to extract themes and details from the transcribed interview material that affect the dissemination, implementation and use of reporting guidelines. The analysis of the content of the interviews will take into consideration the roles,

motivations, context in which the interview was conducted, and identities of the participants and how these details might influence responses.

Thematic analysis will follow a six-phase process described by Braun and Clark[48]:

1. Familiarisation with the content of the entire qualitative dataset. This will occur through TF conducting, transcribing and reading the interviews. Coinvestigators will familiarise themselves with the data by reading the completed transcripts of the interviews. Audio files and electronic copies of the transcriptions of the interviews will be imported into Nvivo 9 as they are completed. Tools within Nvivo 9 will subsequently be used to highlight or select sections, code, compile and retrieve data across interviews.
2. Initial or provisional codes will be developed and data relating to each of them will be collated with Nvivo 9.
3. The initial codes will then be collated into a provisional set of themes with accompanying, relevant data.
4. Potential themes within the data will then be reviewed and checked against respective coded extracts of data and the complete data set.
5. Ongoing refining of themes will occur in order to ensure that they accurately reflect, discriminate between each other and convey the key messages from the interviews.
6. Final refinements to names, definitions and the scope of themes will occur during the writing of a report to be submitted for publication in a peer-reviewed journal. Extracts from the data will be included to illustrate the themes extrapolated from the dataset.

Throughout the analysis, ongoing discussions between the coinvestigators regarding the content of the data will occur. These joint discussions will explore the data, and contribute to the development, testing and refinement of the analytical framework within which codes and themes are developed. It is also expected that the phases of analysis will not progress in a linear process but be iterative or recursive in nature. In other words, analysis will involve moving back and forwards through/to phases in order to refine and test codes and themes identified within the dataset.

The reporting of qualitative research will follow the recommendations in the Consolidated Criteria for Reporting Qualitative Research checklist for interviews.[49]

ETHICS AND DISSEMINATION

The protocol of this study has been assessed and approved by the Research Ethics Committee of the Peninsula College of Medicine and Dentistry, Universities of Exeter and Plymouth (approval number: Jun12/CA/159). All the participants will be asked to complete informed consent forms prior to participating in this study and all the data will be de-identified to ensure anonymity.

Recordings of interviews will be stored securely on the University of Exeter computer network and only the researchers involved in the project will have access to the dataset during the course of the study.

The findings of these studies will show how the implementation and dissemination of the TREND guideline has affected reporting completeness in studies with experimental, non-randomised designs within behavioural and public health research and what factors affect authors' and journal editors' use of reporting guidelines. The results of these studies will be published in peer-reviewed journals and presented at international conferences.

DISCUSSION

These studies aim to add to the existing literature on reporting guidelines by generating evidence of how the implementation and dissemination of TREND has affected reporting completeness in studies with experimental, non-randomised designs within public health. In addition, study 2 represents the first attempt that we are aware of, to collect primary research evidence on perceived barriers and facilitators to the use of reporting guidelines. It is anticipated that the results of the studies will be of interest to an audience well beyond those interested in TREND. Specifically, the findings are likely to be relevant to authors, journal editors, guideline developers and researchers interested in improving the completeness and transparency of reporting of research and improving the effectiveness of dissemination and implementation of reporting guidelines.

Acknowledgements We would like to thank Don Des Jarlais, Iveta Simera and Allison Hirst for comments on an early draft and Bruno da Costa and Kim Eunyoung for their peer review, comments and suggestion on the submitted manuscript. Funding for project including development of this protocol was provided by the National Institute for Health Research (NIHR) Peninsula Collaboration for Leadership in Applied Health Research & Care (PenCLAHRC). The views expressed in this publication are those of the authors and not necessarily those of the NHS, NIHR or the Department of Health, UK.

Contributors TF participated in the study design, writing and editing of the draft manuscript. MP, JP and RA conceived the study, participated in the study design and helped to edit the draft manuscript. All the authors read and approved the final manuscript.

Funding Funding for the project is provided by the Peninsula Collaboration for Leadership in Applied Health Research and Care (PenCLAHRC) from a competitive grant awarded by the National Institute for Health Research (NIHR).

Competing interests None.

Ethics approval Research Ethics Committee of the Peninsula College of Medicine and Dentistry, Universities of Exeter and Plymouth.

Provenance and peer review Not commissioned; externally peer reviewed.

REFERENCES

1. Altman DG, Simera I. Responsible reporting of health research studies: transparent, complete, accurate and timely. *J Antimicrob Chemother* 2010;65:1–3.
2. Deeks JJ, Dinnes J, D'Amico R, *et al*. Evaluating nonrandomised intervention studies. *Health Technol Assess* 2003;7:1–173.

3. Moher D, Weeks L, Ocampo M, *et al*. Describing reporting guidelines for health research: a systematic review. *J Clin Epidemiol* 2011;64:718–42.
4. Begg C, Cho M, Eastwood S, *et al*. Improving the quality of reporting of randomized controlled trials—The CONSORT statement. *JAMA* 1996;276:637–9.
5. Des Jarlais DC, Lyles C, Crepaz N, *et al*. Improving the reporting quality of nonrandomized evaluations of behavioral and public health interventions: the TREND statement. *Am J Public Health* 2004;94:361–6.
6. von Elm E, Altman DG, Egger M, *et al*. Strengthening the reporting of observational studies in epidemiology (STROBE) statement: guidelines for reporting observational studies. *BMJ* 2007;335:806–8.
7. Bossuyt PM, Reitsma JB, Bruns DE, *et al*. Towards complete and accurate reporting of studies of diagnostic accuracy: the STARD initiative. Standards for reporting of diagnostic accuracy. *Clin Chem* 2003;49:1–6.
8. MacPherson H, Altman DG, Hammerschlag R, *et al*. Revised standards for reporting interventions in clinical trials of acupuncture (STRICTA): extending the CONSORT statement. *J Altern Complement Med* 2010;16:ST1–14.
9. Vandenbroucke JP. STREGA, STROBE, STARD, SQUIRE, MOOSE, PRISMA, GNOSIS, TREND, ORION, COREQ, QUOROM, REMARK… and CONSORT: for whom does the guideline toll? *J Clin Epidemiol* 2009;62:594–6.
10. O'Connor A. Reporting guidelines for primary research: Ssaying what you did. *Prev Vet Med* 2010;97:144–9.
11. Eccles M, Foy R, Sales A, *et al*. Implementation Science six years on—our evolving scope and common reasons for rejection without review. *Implement Sci* 2012;7:71.
12. Simera I, Altman DG, Moher D, *et al*. Guidelines for reporting health research: the EQUATOR Network's Survey of Guideline Authors. *PLoS Med* 2008;5:e139.
13. Dechartres A, Charles P, Hopewell S, *et al*. Reviews assessing the quality or the reporting of randomized controlled trials are increasing over time but raised questions about how quality is assessed. *J Clin Epidemiol* 2011;64:136–44.
14. Sanderson S, Tatt ID, Higgins JP. Tools for assessing quality and susceptibility to bias in observational studies in epidemiology: a systematic review and annotated bibliography. *Int J Epidemiol* 2007;36:666–76.
15. Moher D, Jadad AR, Nichol G, *et al*. Assessing the quality of randomized controlled trials: an annotated bibliography of scales and checklists. *Control Clin Trials* 1995;16:62–73.
16. Juni P, Witschi A, Bloch R, *et al*. The hazards of scoring the quality of clinical trials for meta-analysis. *JAMA* 1999;282:1054–60.
17. Jüni P, Altman DG, Egger M. Assessing the quality of controlled clinical trials. *BMJ* 2001;323:42–6.
18. Prady SL, Richmond SJ, Morton VM, *et al*. A Systematic evaluation of the impact of STRICTA and CONSORT recommendations on quality of reporting for acupuncture trials. *PLoS One* 2008;3:e1577.
19. Smidt N, Rutjes AW, van der Windt DA, *et al*. The quality of diagnostic accuracy studies since the STARD statement: has it improved? *Neurology* 2006;67:792–7.
20. Moher D, Jones A, Lepage L. Use of the CONSORT statement and quality of reports of randomized trials: a comparative before-and-after evaluation. *JAMA* 2001;285:1992–5.
21. CONSORT Group. CONSORT Statement Website. http://www.consort-statement.org/about-consort/consort-endorsement/ (accessed 1 Aug 2012).
22. Altman DG. Endorsement of the CONSORT statement by high impact medical journals: survey of instructions for authors. *BMJ* 2005;330:1056–7.
23. Hopewell S, Altman D, Moher D, *et al*. Endorsement of the CONSORT statement by high impact factor medical journals: a survey of journal editors and journal 'Instructions to Authors'. *Trials* 2008;9:20–6.
24. Li X-Q, Tao K-m, Zhou Q-h, *et al*. Endorsement of the CONSORT statement by high-impact medical journals in China: A survey of instructions for authors and published papers. *PLoS One* 2012;7:e30683.
25. Hirst A, Altman DG. Are peer reviewers encouraged to use reporting guidelines? A survey of 116 health research journals. *PLoS ONE* 2012;7:e35621.
26. Plint AC, Moher D, Morrison A, *et al*. Does the CONSORT checklist improve the quality of reports of randomised controlled trials? A systematic review. *Med J Aust* 2006;185:263–7.
27. Turner L, Moher D, Shamseer L, *et al*. The influence of CONSORT on the quality of reporting of randomised controlled trials: an updated review. *Trials* 2011;12:(Suppl 1);A47.
28. Dzewaltowski DA, Estabrooks PA, Klesges LM, *et al*. Trend: An important step, but not enough. *Am J Public Health* 2004;94:1474.
29. Kirkwood B. Making public health interventions more evidence based—TREND statement for non-randomised designs will make a difference. *BMJ* 2004;328:966–7.
30. Armstrong R, Waters E, Moore L, *et al*. Improving the reporting of public health intervention research: advancing TREND and CONSORT. *J Public Health* 2008;30:103–9.
31. Pearson M, Peters J. Outcome reporting bias in evaluations of public health interventions: evidence of impact and the potential role of a study register. *J Epidemiol Community Health* 2012;66:286–9.
32. Greenhalgh T, Robert G, Macfarlane F, *et al*. Diffusion of innovations in service organizations: systematic review and recommendations. *Milbank Q* 2004;82:581–629.
33. Robert G, Greenhalgh T, MacFarlane F, *et al*. Adopting and assimilating new non-pharmaceutical technologies into health care: a systematic review. *J Health Serv Res Policy* 2010;15:243–50.
34. Kulkarni A, Aziz B, Shams I, *et al*. Comparisons of citations in web of science, scopus, and google scholar for articles published in general medical journals. *JAMA* 2009;302:1092–6.
35. Li J, Burnham JF, Lemley T, *et al*. Citation analysis: comparison of Web of Science®, Scopus™, SciFinder®, and Google Scholar. *J Electron Resources Med Libr* 2010;7:196–217.
36. Spring M. *Applying the principles of EBM to public health—searching for public health evidence—the experience at the National Institute for Health and Clinical Excellence (NICE)*. Helsinki, Finland: European Association for Health Informatics and Libraries, 2008.
37. Fraser C, Murray A, Burr J. Identifying observational studies of surgical interventions in MEDLINE and EMBASE. *BMC Med Res Methodol* 2006;6:41.
38. Furlan AD, Irvin E, Bombardier C. Limited search strategies were effective in finding relevant nonrandomized studies. *J Clin Epidemiol* 2006;59:1303–11.
39. Egan M, MacLean A, Sweeting H, *et al*. Comparing the effectiveness of using generic and specific search terms in electronic databases to identify health outcomes for a systematic review: a prospective comparative study of literature search methods. *BMJ Open* 2012;2:e001043.
40. Lachin JM. Introduction to sample size determination and power analysis for clinical trials. *Control Clin Trials* 1981;2:93–113.
41. Davidoff F, Batalden P, Stevens D, *et al*. Publication guidelines for quality improvement in health care: evolution of the SQUIRE project. *Qual Saf Health Care* 2008;17(Suppl 1):i3–9.
42. Thomas BH, Ciliska D, Dobbins M, *et al*. A process for systematically reviewing the literature: providing the research evidence for public health nursing interventions. *Worldviews Evid Based Nurs* 2004;1:176–84.
43. Jackson R, Ameratunga S, Broad J, *et al*. The GATE frame: critical appraisal with pictures. *Evid Based Med* 2006;11:35–8.
44. Wells GA, Shea B, O'Connell D, *et al*. The Newcastle-Ottawa Scale (NOS) for assessing the quality of nonrandomised studies in meta-analyses. http://www.ohri.ca/programs/clinical_epidemiology/oxford.htm(accessed 23 May 2012).
45. National Institute for Health and Clinical Excellence. *Methods for development of NICE public health guidance*. London: National Institute for Health and Clinical Excellence, 2009.
46. Kunath F, Grobe HR, Rucker G, *et al*. Do journals publishing in the field of urology endorse reporting guidelines? A survey of author instructions. *Urol Int* 2012;88:54–9.
47. Kelley K, Clark B, Brown V, *et al*. Good practice in the conduct and reporting of survey research. *Int J Qual Health Care* 2003;15:261–6.
48. Braun V, Clarke V. Using thematic analysis in psychology. *Qual Res Psychol* 2006;3:77–101.
49. Tong A, Sainsbury P, Craig J. Consolidated criteria for reporting qualitative research (COREQ): a 32-item checklist for interviews and focus groups. *Int J Qual Health Care* 2007;19:349–57.

Evaluating the impact and use of Transparent Reporting of Evaluations with Non-randomised Designs (TREND) reporting guidelines

Thomas Fuller, Mark Pearson, Jaime L Peters, et al.

BMJ Open 2012 2:
doi: 10.1136/bmjopen-2012-002073