Cochrane Library

Design-based methods to influence the completeness of response to self-administered questionnaires (Protocol)

Healy P, Edwards PJ, Smith V, Murphy E, Newell J, Burke E, Meskell P, Galvin S, Lynn P, Stovold E, McCarthy B, Biesty LM, Devane D

www.cochranelibrary.com

WILEY

T A B L E   O F   C O N T E N T S

Design-based methods to influence the completeness of response to self-administered questionnaires

Patricia Healy[1], Philip James Edwards[2], Valerie Smith[3], Edel Murphy[4], John Newell[5], Eimear Burke[1], Pauline Meskell[6], Sandra Galvin[1,7], Peter Lynn[8], Elizabeth Stovold[9], Bernard McCarthy[1], Linda M Biesty[1], Declan Devane[1,7]

[1]School of Nursing and Midwifery, National University of Ireland Galway, Galway, Ireland. [2]Faculty of Epidemiology and Population Health, London School of Hygiene & Tropical Medicine, London, UK. [3]School of Nursing and Midwifery, Trinity College Dublin, Dublin, Ireland. [4]PPI Ignite Programme, National University of Ireland Galway, Galway, Ireland. [5]School of Mathematics, Statistics and Applied Mathematics, National University of Ireland, Galway, Ireland. [6]Department of Nursing and Midwifery, University of Limerick, Limerick, Ireland. [7]HRB-Trials Methodology Research Network, National University of Ireland Galway, Galway, Ireland. [8]Institute for Social and Economic Research (ISER), University of Essex, Colchester, UK. [9]Population Health Research Institute, St George's, University of London, London, UK

Contact address: Patricia Healy, School of Nursing and Midwifery, National University of Ireland Galway, University Road, Galway, Ireland. Patricia.healy@nuigalway.ie.

ABSTRACT

This is a protocol for a Cochrane Review (Methodology). The objectives are as follows:

To evaluate the effectiveness of design-based methods to influence the completeness of item response to self-administered questionnaires. This will be achieved by assessing the effects of aspects of style, appearance and layout of self-administered questionnaires on the proportion of items completed in returned questionnaires.

BACKGROUND

Description of the problem or issue

Survey methods using questionnaires comprise a series of standardised questions designed for gathering information about respondents' attributes, behaviours, beliefs, knowledge, attitudes or opinions (Alreck 2004; Rowley 2014). Questionnaires are one of the most frequently used means of collecting data and are used widely in research as they offer one of the least expensive modes of collecting data from relatively large samples (Bowling 2000; Carter 2000). Typically, the questionnaire respondents are a sample drawn from a wider population, and are chosen to represent that population. Questionnaires can be self-administered or interviewer-administered. However, using questionnaires is not without difficulties and they may fail to collect the required data. The absence of an interviewer when using self-administered questionnaires means that they are less susceptible to information bias (e.g. social desirability bias) but are more prone to missing data concerning sensitive or financial information (Bowling 2005). The

issues of data quality and missing data from questionnaires can pose serious problems for researchers. The validity and reliability of the findings of research studies are determined by the quality of the data collected. Missing data occurs in situations where the whole questionnaire is not returned; this is referred to as unit nonresponse, or where items are not completed in the returned questionnaire; this is referred to as item nonresponse. Both unit and item nonresponse can result in bias and reduced statistical power in a study.

Description of the methods being investigated

There is a substantial body of empirical evidence around interventions aimed at maximising the return of questionnaires that have been distributed for research purposes (Edwards 2007; Edwards 2009; McColl 2001; Nakash 2006). Numerous strategies, both incentive-based (e.g. cash incentive, gift card) and design-based (e.g. shorter length, booklet format), have been devised to increase unit response rates to questionnaires. The successful return of self-administered questionnaires, however, does not ensure that responses have been provided to all of the items in the questionnaire. In addition to the successful return of the questionnaire (unit response rates), researchers ought also to be concerned about the completeness of responses to the items in those questionnaires (item response rates). The bias introduced by item nonresponse depends on both the item nonresponse rate and the true distribution of the missing values. In addition to the impact of missing data mechanisms and missing data patterns on research results, it must also be acknowledged that the proportion of missing data items affects the overall data quality. Although, the literature does not reflect agreement on a minimum acceptable percentage of missing data in a data set for valid statistical inferences (Dong 2013), a number of suggestions have been made. The American Association for Public Opinion Research, as a general rule of thumb, suggest that if less than 50% of all essential or crucial questions in the survey are answered it is incomplete, 50% to 99% answered equals partially complete, and 100% equals complete (AAPOR 2011). Though knowledge of the rate alone is not informative regarding the extent of bias, it is clear that higher item nonresponse rates have the potential to be associated with greater bias. Although there are statistical approaches to managing missing data such as imputation techniques and pairwise deletion, it is possible that there are systematic differences between respondents who complete certain questionnaire items and respondents who do not (De Leeuw 2001). Thus, data obtained may not be representative of the sample. Reduction of item nonresponse through better questionnaire design would reduce the need for methods for managing missing data and minimise associated bias. Various design-based issues that may influence the extent to which items in a questionnaire are completed fully have been described to some extent in the literature. These include print format (single-sided versus double-sided), order of questions, open-ended versus closed-ended questions, length of questionnaire, ease of response format, sensitivity of the question topic, salience of the question and the layout and general appearance of the questionnaire (Boynton 2004; Dillman 2000; Dillman 2008; Edwards 2010; Fowler 2008; Jenkins 1995; McColl 2001; Rowley 2014).

How these methods might work

It is recognised that the responses given to self-administered survey questions are the result of a complex interaction between the person completing the questions, the mode of delivery of the questionnaire and the questionnaire design (Dillman 2008; Lynn 2008; Tourangeau 2000; Tourangeau 2004). However, the traditional good practice principles of questionnaire design have a limited empirical basis. A number of authors have previously recommended the need for further studies of methods that might improve the quality and quantity of the data collected by questionnaires (Cavusgil 1998; De Leeuw 2001; Edwards 2009; Edwards 2010; Jenkins 1995; Wilks 2007), but a Cochrane Review has not been undertaken on the subject. This review will fill that gap. It will evaluate the effectiveness of design-based methods to influence item response in self-administered questionnaires and is complimentary to the evidence relating to unit response.

Why it is important to do this review

Questionnaires are used widely for research as they are an economic and pragmatic way to collect large volumes of data. Pen and paper questionnaires remain an important method of data collection in epidemiological investigations. In a review of over 2000 analytic epidemiological research articles published in high-impact medical and epidemiological journals during 2008 and 2009, more than one quarter relied on pen and paper questionnaires as their mode of data collection (van Gelder 2010). In addition to the return of the questionnaire, successful data collection by questionnaire depends on the participant completing the items in the questionnaire that collect the required data. Following a review of unit response rates to questionnaires, Edwards 2010 suggests that further research is needed into the types of questions and the style, appearance and layout of questionnaires that are effective in increasing data quality and completeness. This systematic review will synthesise the effectiveness of these design-based measures for influencing item response during completion of self-administered questionnaires. Finding ways to maximise item response in studies that collect data by questionnaire could improve data completeness, minimise bias, improve the validity of study findings and limit waste of time and financial resources. The potential beneficiaries of this review are wide ranging as the questionnaire remains a widely used data collection instrument across many diverse areas of research.

OBJECTIVES

To evaluate the effectiveness of design-based methods to influence the completeness of item response to self-administered questionnaires. This will be achieved by assessing the effects of aspects of style, appearance and layout of self-administered questionnaires on the proportion of items completed in returned questionnaires.

METHODS

Criteria for considering studies for this review

Types of studies

We will include randomised trials comparing at least one design-based intervention intended to increase item response to self-administered questionnaires. This will include studies of interventions to increase total unit response as some of those interventions may inevitably increase item response. We will also include randomised studies within trials (SWAT) (Clarke 2015), where we will extract data relevant to the design intervention rather than the main host study. We will exclude quasi-randomised trials. Given the review objective, we do not expect to find cluster or cross-over trials.

Types of data

We will include randomised trials collecting data by self-administered questionnaire. We are interested in questionnaires that are designed to be completed without any direct interaction with the researcher. For the purposes of this review, the term self-administered questionnaire is defined to mean structured surveys used to elicit predominantly quantitative information, by means of direct questions, from informants by self-completion (McColl 2001; page 4) using the traditional "pencil-and-paper" methods of recording responses.

Types of methods

We will consider studies that describe any design-based method applied to a self-administered questionnaire to influence item nonresponse in the returned questionnaire. Design-based methods may include aspects of style, appearance and layout of the questionnaire such as questionnaire length, the response format or inclusion of sensitive questions. Questionnaires sent to participants by post or handed to them in person but subsequently self-completed will be included, but those completed during telephone or face to-face interviews or online will be excluded. We have excluded online questionnaires because online mode of administration can include

options to force item completion not available with traditional "pencil-and-paper".

Types of outcome measures

Primary outcomes

The number of items completed in returned questionnaires as a proportion of all items in the questionnaire that should have a response. In cases where values are missing for obvious reasons, such as legitimate skip items, these are not considered nonresponses.

Secondary outcomes

- The proportion of returned questionnaires where responses have been given to all items. This will be represented by an item nonresponse rate of 0%
- The proportion of returned questionnaires where responses have been given to 90% of all items. This will be represented by an item nonresponse rate of 10%
- The proportion of returned questionnaires where responses have been given to 80% of all items. This will be represented by an item nonresponse rate of 20%
- The proportion of returned questionnaires where responses have been given to 50% of all items. This will be represented by an item nonresponse rate of 50%

These outcomes may not be available for all studies, but will be measured where available. Other outcomes not reported in the protocol whose importance is realised after the protocol is written or when the analysis is done may be added but will be identified clearly as post hoc.

Search methods for identification of studies

Search strategies are developed to achieve a balance between sensitivity and specificity. The search strategy will be modified as necessary for use with multiple databases to ensure that the search is comprehensive, thorough and objective. Restricting search terms to the title and abstract field only, by using permutations of subject term combinations, or by using fewer search terms will increase the specificity of the searches.

Electronic searches

We will search MEDLINE using the search strategy outlined in Appendix 1. We will adapt this search strategy for use with other databases including:
- Embase (via Ovid)
- PsycINFO (via Ovid)
- CINAHL Plus (via EBSCO)
- MEDLINE (via Ovid)

- Educational Resources Information Centre (ERIC)
- British Education Index
- Sociological Abstracts
- Social Science Citation Index
- Science Citation Index
- Cochrane Methodology Register
- Cochrane Central Register of Controlled Trials
- Open Grey (http://www.greynet.org)
- ProQuest Dissertation and Thesis
- Index to Theses (Great Britain and Ireland) (http://www.theses.com/)

This list may be adjusted after some sample searching using the search strategy shown in Appendix 1 (modified as appropriate).

Searching other resources

We will also search international registers of current and ongoing clinical trials including the ClinicalTrials.gov and the WHO International Clinical Trials Registry Platform (ICTRP) portal (Ghersi 2009). Language restrictions will not be applied to the search. We will check the reference lists of relevant included studies and systematic reviews identified through the electronic searches for additional references. If necessary, we will contact authors of ongoing trials or relevant publications in press for additional information on relevant studies.

Data collection and analysis

Selection of studies

Two review authors will independently screen titles and abstracts of citations retrieved by searching against a pre-specified eligibility criteria based on types of studies (randomised trials), types of interventions (design-based measures), participants (self-administered questionnaires) and measured outcomes (item completion/non-response). The records will be sorted into the following groups; 'include', 'exclude' or uncertain. Studies for which there is uncertainty will have their full-text papers reviewed by both review authors to reduce the potential for random errors and bias. If, after discussion, there is still disagreement regarding study inclusion, a third review author will review the full paper and consider its eligibility for inclusion. We will import the references of potentially eligible studies into EndNote and remove duplicate records of the same reports. Each of the full--text reports will be obtained and assessed by two review authors to determine if they meet the inclusion criteria for the review and any disagreement on the eligibility of included studies will be resolved through discussion. Where resolution is not possible, we will discuss issues raised with a third review author. In addition, we will contact study authors in order to identify unavailable/unclear data.

Data extraction and management

Two review authors will independently extract data from each study using tailored data abstraction forms that will be piloted and improved as necessary. Any discrepancies will be discussed and where resolution is not possible a third review author will be consulted. We will extract the following data.

- Author details
- Publication year of study
- Data source (journal of publication, other)
- Language
- Setting
- Country
- Study methods including study design
- Study participants, numbers and proportions in each intervention group
- Description of self-administered survey questionnaires used
- Intervention: e.g. number of pages in the questionnaire; the inclusion of sensitive questions, the layout of the questionnaire, the questionnaire topic (healthcare/non healthcare) etc.
- Comparison: details of comparison, e.g. shorter questionnaire, different size font, use of filter questions etc.
- Data to assess the risk of bias of included studies e.g. sequence generation, allocation concealment, blinding of study participants and personnel, blinding of outcome assessors, withdrawals or incomplete outcome data, selective reporting or other sources of bias
- Outcomes: review pre-specified outcomes

Assessment of risk of bias in included studies

We will consider aspects of the design, conduct, analysis and reporting of the study that could cause the effect of an intervention to be underestimated or overestimated and thereby affect the internal or external validity of the results. Two review authors (PH and DD or EM) will assess the risk of bias for each included study independently using Cochrane's criteria for assessing the risk of bias in randomised trials (Higgins 2011). We will contact study authors when necessary information or data are not available in the published reports or if clarification is required. Two review authors will apply the 'Risk of bias' criteria to each study independently and differences will be resolved by consulting a third review author (VS) if necessary. We will assess the risk of bias across the following domains.

Selection bias: random sequence generation

Selection bias occurs when the groups formed for comparison have not been created through random allocation and are different in some way that may affect outcome (Torgerson 2003). The rules for allocating interventions to participants in the studies will be reported so that we can identify whether there is a risk that groups assigned to different arms may not have been comparable. We will base our judgements on the following criteria:

- if sequence generation is truly random (e.g. computer-generated random assignment), studies will be deemed at low risk;
- if sequence generation is not specified and we are unable to obtain relevant information from study authors, the study will be considered as an unclear risk;
- if there is a quasi-random sequence generation e.g. alternation: the study will be excluded (see Types of studies);
- if sequence generation uses any non-random process (e.g. odd or even date of birth; hospital or clinic record number), the study will be considered at high risk;
- If there is an unusually large number of differences between intervention group sizes and/or baseline characteristics, the study will be considered at high risk.

Selection bias: allocation concealment

Allocation concealment refers to the methods used by a study to ensure that researchers and participants cannot foresee treatment assignments (Nelson 2014). We will report the process for allocation concealment used in the studies so that we can identify if appropriate steps were taken to ensure that knowledge of the allocation sequence was not possible before the assignment of interventions to participants. We will base our judgements on the following criteria:
- If the study used opaque, sequentially-numbered sealed envelopes or centralised, off-site allocation by a third party, studies will be deemed at low risk;
- If the allocation concealment is not specified and we are unable to ascertain whether the allocation concealment was protected before and until assignment, the study will be considered as an unclear risk;
- If the studies have inadequacies in their allocation concealment, e.g. if non-opaque envelopes, unsealed envelopes, self-selection or clinician-selection, the study will be considered at high risk.

Performance bias

A. Blinding of participants

Performance bias refers to bias related to differential provision of care and follow-up, other than the interventions of interest, due to knowledge of the intervention received (Nelson 2014). The process for blinding in the studies will be reported so that we can identify if appropriate steps were taken to ensure that knowledge of the allocation of intervention to participants was not possible. We will base our judgements on the following criteria:
- If the study participants were unaware whether they received the intervention or control, or if we judge that the lack of blinding would be unlikely to affect results, the study will be deemed at low risk;

- If the blinding of study participants was not specified and we are unable to ascertain whether performance bias is a risk, the study will be considered as an unclear risk;
- If it was not possible to blind participants to the intervention to which they have been assigned and that lack of blinding would be likely to affect results, the study will be considered at high risk for performance bias.

B. Blinding of personnel

We will base our judgements on the following criteria:
- If the study personnel were unaware whether the groups were intervention or control or if we judge that the lack of blinding would be unlikely to affect results, the study will be deemed at low risk;
- If the blinding of study personnel was not specified and we are unable to ascertain whether performance bias is a risk, the study will be considered as an unclear risk;
- If it was not possible to blind personnel to the intervention to which participants have been assigned and that lack of blinding would be likely to affect results, the study will be considered at high risk for performance bias.

Given the nature of the intervention under review, we do not expect that blinding of participants or personnel will have been likely.

Detection bias: blinding of outcome assessors

Detection bias refers to bias related to whether the outcome assessor was blinded to group allocation. We expect it to be likely that it will not be possible to blind outcome assessors to the design differences between control and intervention questionnaires. In addition, we expect that the assessment of our outcomes is unlikely to be influenced by knowledge of the intervention received (i.e. if an outcome is unaffected by blinding) and therefore will not assess detection bias.

Attrition bias: incomplete outcome data

Attrition bias refers to bias related to missing data or loss to follow-up/withdrawals, where participants lost to follow-up differ systematically from those who remain in the trial (Nelson 2014). We will explore withdrawals or incomplete outcome data due to exclusions or attrition (the number randomised minus any participants whose questionnaires are known to be missing) so that we can identify the extent of attrition bias. Although the literature is ambiguous on a minimum acceptable unit response rate, there is general consensus that at least half of the sample should have completed the survey instrument (Draugalis 2008). Therefore, we will base our judgements on the following pragmatic criteria:

- if less than 50% of the questionnaires (unit response)are missing and are spread equally across groups, the study will be deemed at low risk;
- if the percentage of missing questionnaires or the spread of missing questionnaires is not clear, the study will be deemed at unclear risk,
- if 50% or more of the questionnaires are missing, if the missing questionnaires are not equally spread across groups or if the missing questionnaires were not handled appropriately (intention-to-treat analysis, imputation), the study will be deemed at high risk.

Selective reporting bias

Selective reporting bias refers to bias due to a tendency to under-report results based on the direction or statistical significance of those results (Kirkham 2010; Nelson 2014). We will explore whether all pre-specified primary and important secondary outcomes mentioned in the protocol and methodology sections of the studies are reported in results sections. We will base our judgements on the following criteria:
- if all outcomes are both listed in the protocol and methodology and then reported in the results:,the study will be deemed at low risk;
- if we cannot ascertain from the information provided by study authors, the study will be deemed at unclear risk;
- if all outcomes in the protocol and methodology are not reported in the results or if outcomes reported in the results were not listed in the protocol and methodology, the study will be deemed at high risk;
- if outcomes are only partly reported in the results or if an obvious outcome is not mentioned in the study, the study will be deemed at high risk.

Other potential sources of bias

We will assess the studies for other potential biases (e.g. recruitment bias: imbalance in respondent characteristics) using the following criteria:
- If there is no evidence of other sources of bias, the study will be deemed at low risk;
- If there is incomplete information regarding a problem which may lead to bias, the study will be deemed at unclear risk;
- If there is one or more important risks of bias e.g. flawed study design, the study will be deemed at high risk.

We will summarise the information extracted in the 'Characteristics of included studies' table. We anticipate that information may not be available in all studies, particularly studies outside health care. The information will be sought from authors if unclear from the published study data. For each included study, review authors will classify each domain as presenting low, high, or unclear risk of bias. Any discrepancies between the two review authors conducting the assessment of risk of bias will be resolved through discussion. If no agreement can be reached, a third review author (VS) will act as an arbiter.

Measures of the effect of the methods

Effects of intervention for dichotomous outcome data will be determined using a risk ratio (RR) with a 95% confidence intervals (CIs).

When interventions are evaluated at more than two levels (e.g. short, long, very long questionnaire), we will combine levels to create a dichotomy. Ordinal scale data outcomes reported will be collapsed into dichotomous outcomes.

For continuous data, we will calculate the mean difference (MD) and 95% CIs if the measurement scale is the same. If the scale is different, we will use standardised mean differences (SMD) with 95% CIs.

When data to calculate standard deviations (SDs) are missing from studies, and it is not possible to obtain the result from study authors, we will use the mean value for the SD of other included studies that reported that outcome.

Where continuous outcome data are reported as medians and Interquartile ranges/ranges instead of means and SDs; this will be reported narratively.

Unit of analysis issues

The unit of analysis will be the individual survey/questionnaire (the unit). We will group trials according to the type of intervention (questionnaire length, format, layout etc.) where the interventions are similar in form and content. We do not expect to identify any cluster-randomised trials.

Dealing with missing data

For included studies, we will note the level of attrition. Participants will be analysed according to the arm to which they were randomised, even if they do not receive the allocated intervention.

Assessment of heterogeneity

We will assess heterogeneity visually through inspection of forest plots. We will assess statistical heterogeneity in each meta-analysis using the Chi^2 test for heterogeneity and we will quantify the degree of heterogeneity observed in the results using the I^2 and Tau^2 statistic (Higgins 2011). We will regard heterogeneity as substantial if an I^2 is greater than 30% and either the Tau^2 is greater than zero, or there is a low P value (< 0.10) in the Chi^2 test for heterogeneity. If we identify substantial heterogeneity (> 30%), we plan to explore it by pre-specified subgroup analysis.

Assessment of reporting biases

We will conduct a comprehensive search of multiple bibliographic databases and trial registries in order to minimise the risk of publication bias which can arise when the dissemination of research findings is influenced by the nature and direction of results. Searches will be conducted without language restrictions. Duplicate publications will be identified. If 10 or more studies are included in a meta-analysis, we will create a funnel plot of the intervention-effect estimates against a measure of the studies size or precision to investigate whether bias may exist. We will use the funnel plot test proposed by Egger 1997. If we notice asymmetry we cannot conclude that reporting biases exist however. We will consider other possible sources of asymmetry such as the sample sizes, methodological design and presence and possible influence of outliers and subsequently perform a sensitivity analysis.

Data synthesis

We will analyse our data using RevMan (RevMan 2014). Our intention is to calculate effect estimates using an intention-to-treat analysis, but we expect that there will be some participants for whom outcome data (item-response rates) are unavailable and these will be excluded from the analyses. We will assess the clinical and methodological diversity between included studies qualitatively. We expect that the studies we will be including in this review will vary in terms of their sample characteristics, interventions tested and comparisons applied and therefore we will use a random-effects model to incorporate heterogeneity among studies. Random-effects models are based on the assumption that the true effect might vary across samples and studies. Random-effects meta-analysis can incorporate heterogeneity into meta-analysis but does not fix it. For each outcome reported, we will present the random-effects estimate with its 95% confidence interval, and the estimates of Tau² and I². In the absence of sufficient homogeneity, we will present the quantitative results in a tabular form and describe them narratively. Details of each intervention will be presented in a table of study characteristics.

Subgroup analysis and investigation of heterogeneity

If there is evidence of statistical heterogeneity among the trials, we will explore using subgroup and/or sensitivity analyses to identify the causes of heterogeneity. Subgroup analyses involve dividing the studies into subgroups of those with similar characteristics (e.g. intervention type) and performing separate meta-analyses for each group of potentially homogeneous studies. This test provides an effect estimate within subgroups and a significance test for that estimate. Trials will be grouped according to the type of design-based intervention evaluated (e.g. questionnaire length) and interventions will be grouped when they are similar in form and content. Intervention categories may include, but are not limited to the following design-based features.

- Questionnaire length (long versus short)
- Questionnaire format (booklet versus stapled pages)
- Questionnaire appearance (coloured versus white)
- Questionnaire lay-out (horizontal versus vertical orientation)
- Print format (single versus double-sided)

We will assess subgroup differences by interaction tests available within RevMan (RevMan 2014). We will report the results of subgroup analyses quoting the Chi² statistic and P value, and the interaction test I² value.

Sensitivity analysis

Sensitivity analysis tests the impact of decisions that were made during the review process to determine whether results are robust (consistent) under different assumptions. Different subgroups of studies are synthesised while systematically excluding some studies to determine how this affects the review conclusions. For example, studies below a certain quality threshold may be excluded and then the intervention effect is recalculated to examine the impact of that study on the overall results. Sensitivity analysis can also determine whether results were robust across different methods of handling missing data. We will conduct a sensitivity analysis based on trial quality assessed by concealment of allocation, high attrition rates, or both, with poor-quality studies being excluded from the analyses in order to assess whether this makes any difference to the overall result.

ACKNOWLEDGEMENTS

REFERENCES

Additional references

AAPOR 2011

AAPOR:The American Association for Public Opinion Research. *The American Association for Public Opinion Research Standard definitions: Final dispositions of case codes and outcome rates for surveys*. Lenexa, Kansas: The American Association for Public Opinion Research (www.AAPOR.org), 2011.

Alreck 2004

Alreck P, Settle R. *The Survey Research Handbook*. 3rd Edition. New York: McGraw-Hill, 2004.

Bowling 2000

Bowling A. *Research Methods in Health: Investigating Health and Hhealth Services*. Buckingham: Open University Press, 2000.

Bowling 2005

Bowling A. Mode of questionnaire administration can have serious effects on data quality. *Journal of Public Health* 2005;**27**(3):281-91.

Boynton 2004

Boynton P, Greenlagh T. Selecting, designing, and developing your questionnaire. *BMJ* 2004;**328**(7451): 1312-5.

Carter 2000

Carter Y, Shaw S, Thomas C. *The Use and Design of Questionnaires*. London: The Royal College of General Practitioners, 2000.

Cavusgil 1998

Cavusgil S, Elvey-Kirk L. Mail survey response behaviour. *European Journal of Marketing* 1998;**32**(11/12):1165-92.

Clarke 2015

Clarke M, Savage G, Maguire L, McAneney H. The SWAT (study within a trial) programme; embedding trials to improve the methodological design and conduct of future research. *Trials* 2015;**16**(Suppl. 2):P209. DOI: 10.1186/1745-6215-16-S2-P209

De Leeuw 2001

De Leeuw E. Reducing missing data in surveys: an overview of methods. *Quality and Quantity* 2001;**35**(2):147–60.

Dillman 2000

Dillman D. *Mail and internet surveys: The tailored design method*. New York: Wiley, 2000.

Dillman 2008

Dillman D. The logic and psychology of constructing questionnaires. In: de Leeuw E, Hox J, & Dillman D editor (s). *International Handbook of Survey Methodology*. New York: Taylor & Francis Group, 2008:161-75.

Dong 2013

Dong Y, Peng CY. Principled missing data methods for researchers. *SpringerPlus* 2013;**2**:222–39.

Draugalis 2008

Draugalis J, Coons S, Plaza C. Best practices for survey research reports: A synopsis for authors and reviewers. *American Journal of Pharmaceutical Education* 2008;**72**(1): 11. [PUBMED: 18322573]

Edwards 2009

Edwards PJ, Roberts I, Clarke MJ, DiGuiseppi C, Wentz R, Kwan I, et al. Methods to increase response to postal and electronic questionnaires. *Cochrane Database of Systematic Reviews* 2009, Issue 3. DOI: 10.1002/14651858.MR000008.pub4

Edwards 2010

Edwards P. Questionnaires in clinical trials: guidelines for optimal design and administration. *Trials* 2010;**11**:2. DOI: https://doi.org/10.1186/1745-6215-11-2

Egger 1997

Egger M, Smith GD, Schneider M, Minder C. Bias in meta-analysis detected by a simple, graphical test. *BMJ* 1997; **315**(7109):629–34. DOI: http://dx.doi.org/10.1136/bmj.315.7109.629

Fowler 2008

Fowler F, Cosenza C. Writing effective questions. In: de Leeuw, E. Hox, J, DillmanD editor(s). *International Handbook of Survey Methodology*. New York: Taylor & Francis Group, 2008:161-75.

Ghersi 2009

Ghersi D, Pang T. From Mexico to Mali: four years in the history of clinical trial registration. *Journal of Evidence-based Medicine* 2009;**2**(1):1–7. DOI: 10.1111/j.1756-5391.2009.01014.x; PUBMED: 21348976

Higgins 2011

Higgins JPT, Green S, editor(s). Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0 (updated March 2011). The Cochrane Collaboration, 2011. Available from handbook.cochrane.org 2011.

Jenkins 1995

Jenkins CR, Dillman DA. Towards a theory of self-administered questionnaire design. *Survey Measurement and Process Quality*. New York: Wiley-Interscience, 1995.

Kirkham 2010

Kirkham JJ, Altman DG, Williamson PR. Bias due to changes in specified outcomes during the systematic review process. *PLOS One* 2010;**5**(3):1–5. DOI: https://doi.org/10.1371/journal.pone.0009810

Lynn 2008

Lynn P. The Problem of nonresponse. In: de Leeuw E, Hox J, & Dillman D editor(s). *International Handbook of Survey Methodology*. New York: Taylor & Francis Group, 2008:35-55.

McColl 2001

McColl E, Jacoby A, Thomas L, Soutter J, Bamford C, Steen N, et al. Design and use of questionnaires: a review of best practice applicable to surveys of health service staff and patients. *Health Technology Assessment* 2001;**5**(31):1–266. DOI: https://dx.doi.org/10.3310/hta5310

Nakash 2006

Nakash R, Hutton J, Jørstad-Stein E, Gates S, Lamb S. Maximising response to postal questionnaires - A systematic review of randomised trials in health research. *BMC Medical Research Methodology* 2006;**6**:5. DOI: https://doi.org/10.1186/1471-2288-6-5

Nelson 2014

Nelson H. *Systematic Reviews to Answer Health Care Questions*. Philadelphia: Lippincott, Williams & Wilkins, 2014.

RevMan 2014 [Computer program]

RevMan 2014. The Cochrane Collaboration. Review Manager (RevMan). Version 5.3. Copenhagen: The Nordic Cochrane Centre, The Cochrane Collaboration, 2014.

Rowley 2014

Rowley J. Designing and using research questionnaires. *Management Research Review* 2014;**37**(3):308–30.

Torgerson 2003

Torgerson C. *Systematic Reviews: the Continuum Rresearch Methods Series*. London: Continuum International Publishing Group, 2003.

Tourangeau 2000

Tourangeau R, Rips L, Rasinski K. *The Psychology of SurveyRresponse*. Cambridge: Cambridge University Press, 2000.

Tourangeau 2004

Tourangeau R, Couper M, Conrad F. Spacing, position, and order: interpretive heuristics for visual features of survey questions. *Public Opinion Quarterly* 2004;**68**(3):368–93.

van Gelder 2010

van Gelder M, Bretveld R, Roeleveld N. Web-based questionnaires: the future ineEpidemiology?. *American Journal of Epidemiology* 2010;**172**(11):1292-8.

Wilks 2007

Wilks R, Younger N, Mullings J, Zohoori N, Figueroa P, Tulloch-Reid M, et al. Factors affecting study efficiency and item nonresponse in health surveys in developing countries: the Jamaica national healthy lifestyle survey. *BMC Medical Research Methodology* 2007;**7**:13. DOI: 10.1186/1471-2288-7-13

* *Indicates the major publication for the study*

A P P E N D I C E S

Appendix 1. Example search strategy (for MEDLINE (Ovid))

1. "Surveys and Questionnaires"/
2. questionnaire$.ti,ab.
3. survey$.ti,ab.
4. (instrument or instruments).ti,ab.
5. or/1-4
6. (self-administ$ or self administ$).ti,ab.
7. self-assessment/
8. (self-assess$ or "self assess$").ti,ab.
9. (self-complet$ or "self complet$").ti,ab.
10. (self-report$ or "self report$").ti,ab.

11. (self-direct$ or "self direct$").ti,ab.
12. or/6-11
13. 5 and 12
14. (item or items).ti,ab.
15. (question or questions).ti,ab.
16. (answer or answers).ti,ab.
17. (response or responses).ti,ab.
18. or/14-17
19. (nonrespond* or non-respond*).ti,ab.
20. (nonrespons* or non-respons*).ti,ab.
21. (miss or missing or missed).ti,ab.
22. (omission or omit*).ti,ab.
23. complete$.ti,ab.
24. bias$.ti,ab.
25. accuracy.ti,ab.
26. incorrect$.ti,ab.
27. (valid$ or invalid$).ti,ab.
28. unanswered.ti,ab.
29. or/19-28
30. (controlled clinical trial or randomized controlled trial).pt
31. (randomized or randomised).ab,ti.
32. placebo.ab,ti.
33. randomly.ab,ti.
34. trial.ab,ti.

(Continued)

35. groups.ab,ti.
36. or/30-35
37. Animals/ not Humans/
38. 36 not 37
39. 13 and 18 and 29 and 38

CONTRIBUTIONS OF AUTHORS

DD, PH and PE devised the study. PH drafted the protocol and all authors provided feedback and approved the final version. PH and ES developed the search strategy and will conduct the search. PH, EB, PM, SG,EM, DD, LB and BMcC will conduct the screening of title and abstracts and full-text papers. DD, PH and EM will select the studies, assess risk of bias and certainty of the evidence. All authors will contribute expertise as required.

DECLARATIONS OF INTEREST

The authors declare no financial conflict of interest.

SOURCES OF SUPPORT

Internal sources

- No sources of support supplied

External sources

- Health Research Board, Ireland.

Patricia Healy is the recipient of a Cochrane Fellowship from the Health Research Board of Ireland under grant number CTF-2015-1590. The funders had no role in the development of this protocol.