Title: Searching for Trial Protocols: a comparison of methods

Authors: Anthea Sutton, Maria Carmen Galvan De La Cruz, Joanna Leaviss, Andrew Booth

Abstract

Introduction: Registration and publication of trial protocols has become increasingly important and a requirement in some sources of funding and publication. Increased access to protocols yields many potential benefits, but there are issues regarding identification of published protocols. The aim of this investigation is to compare methods of retrieval for identifying trial protocols in a systematic review.

Methods: Six stages of searching (checking published trial reports, searching journal websites, Internet searching, trial registers, bibliographic databases, contact with authors) were completed to identify 74 trial protocols.

Results: 57% of the trial protocols were identified upon completion of all six stages of searching. The most comprehensive method was searching trial registers which identified 51% of the protocols. Contact with authors was most effective at uniquely identifying protocols, 12% were retrieved via this single method. Contact with authors was the only effective method of identifying protocols for trials pre-2005.

Discussion: When attempting to identify trial protocols to include in systematic reviews, some methods are relatively quick to undertake but deliver a low yield. The most effective search strategy for most sources was retrieval by trial registration number where available.

Conclusions: For protocols of trial results published pre-2005, review authors should contact authors as a priority. For protocols post-2005, they should check the trial publication for protocol details, search trial registers, contact authors, ceasing searching once a predetermined point of diminishing returns has been reached.

Keywords: Clinical Protocols; Clinical Trials; Information Storage and Retrieval; Systematic Reviews;

****

1. Introduction

Recent years have seen increasing interest in the registration of trial protocols. Beyond the immediate value of the protocol to ensuring implementation fidelity by study collaborators, and its use by those involved in funding, scientific and ethical review (Chan et al, 2013), there is increased recognition of its value to the wider research community. Drivers for this trend include (1) The existence of a protocol, prior to conduct of the trial itself, acts as a deterrent to non-publication of the trial results under circumstances where the trial shows a statistically significant negative outcome (Odutayo et al,2015), points in a direction that the trial sponsor may wish to suppress (Abaid et al, 2007; Dwan et al, 2008) or documents

serious adverse effects (Hartung et al, 2014); (2) Publication of the protocol protects against selective reporting of trial outcomes, allowing retrospective comparison of trial reports against the pre-existing protocol (Chan et al, 2004a; 2004b; Chan & Altman, 2005; Chan et al, 2008; Bourgeois et al, 2010; Dwan et al, 2011); (3) Publication of the protocol also mitigates against potential research waste in avoiding needless duplication of research when a conclusive result is already available (Chalmers & Glasziou, 2009; Glasziou et al, 2014)). In the context of systematic reviews a further reason for protocol registration lies in (4) facilitating the identification of studies, whether in progress or already completed (Pandis et al, 2016), as a starting point for tracing subsequent trial reports for potential inclusion in a review. Finally, public documentation of trial protocols may promote trial awareness for the public and other investigators (Frederickson & Ifeld, 2011). Such multifarious use has strengthened the need for consensual guidance to improve the content and quality of trial protocols (Tetzlaff et al, 2012; Chan et al, 2013a; 2013b) and to standardise the type of data recorded in trials registers (Nurbhai et al, 2005).

In each of the above circumstances the utility of the protocol is determined by its accessibility (Chan, 2012; Viergever & Li, 2015); both in identifying that a protocol exists and, subsequently, in being able to obtain the most current document of record (Getz et al, 2011; Huic et al, 2011). Potential sources of protocols, or information required to identify protocols (Chan, 2012; Wolfe et al, 2013), include trial registers (Tai et al, 2012), results databases (Zarin et al, 2011), regulatory agency submissions and databases, contact with trialists and/or sponsors (Ko et al, 2011; Schroll et al, 2013), litigation documents, conference abstracts (Huynh et al, 2011; Scherer et al, 2012), and general Internet searches. In their study of protocols for trials subsequently published in the Lancet, Al-Mazouki and colleagues (2008) reported identifying protocols for only 37 of the 64 trials (50 reports) through database searching and contact with authors. Contact with authors is best effected by email, preferably including a request for details of a specific named trial (Young & Hopewell, 2011). Trials registers are noted as an important source for identifying additional RCTs in systematic reviews (Baudard et al, 2017). Searching trials registers is mandatory for Cochrane Reviews (Lefebvre et al, 2013) but their currency (Law et al, 2011; Viergever & Ghersi, 2011; Gill, 2012) and individual and collective coverage (Dwan et al, 2011; Viergever & Ghersi, 2011) remains uneven, search functionality may be limited (Glanville et al, 2014) and their indexing may be deficient.

One potentially key issue is the strength of “linkage” between individual trial registries and published trial results (Bashir & Dunn, 2016). Strong bi-directional linkage is critical with protocols offering a supplementary identification route for trials and trials, in turn, requiring links to protocols in order to assess potential outcome reporting bias. Trial registration numbers may assist identification of a protocol for a known trial (van de Wetering et al, 2012) but do nothing to assist retrieval of previously unknown trials that match the inclusion criteria

for a systematic review. The plethora of terms and synonyms required to achieve sensitive retrieval of such trials from a trial register may prove prohibitive given registers' limited functionality when compared to bibliographic databases (Glanville et al, 2014).
Empirical studies are required to establish (i) what proportion of trial protocols are retrievable using pragmatic methods of retrieval; (ii) what are the most effective methods for retrieving protocols; (iii) what are the main barriers in retrieving protocols; and (iv) whether the most easily retrieved copy of the protocol remains the document of record. Such studies are time consuming and labour intensive and so, realistically, are best undertaken on a case study basis, additional to the requirements to populate a funded systematic review.

The present study was conducted as part of the process of quality assessment of studies included in a systematic review. The review aimed to evaluate the effectiveness of behavioural modification interventions in improving the physical symptoms and functioning of individuals with medically unexplained symptoms (MUS). This complex review exhibited heterogeneity across multiple areas. The included population was 'MUS'. This term may be used to refer to patients with symptoms that persist over many weeks, but that cannot easily be explained even after numerous physical examinations and tests. These symptoms may also be referred to as 'functional' symptoms and described under 'functional somatic syndromes' such as irritable bowel syndrome, fibromyalgia or chronic fatigue syndrome. The review sought to evaluate a broad and varied range of interventions all aiming to modify patient behaviour, most commonly with cognitive therapies, behavioural therapies, or exercise interventions. The number and nature of outcomes reported in the trials was also widely varied. Due to population differences, a range of physical symptoms were recorded, with physical and emotional functioning reported but measured with diverse instruments. Published reports of studies are often restricted by word limits, and it is useful to refer directly to protocols for clarity on these aspects of studies.

Within the context of the funded systematic review, an investigation was conducted to identify the 74 trial protocols relating to the potentially included studies. The aim of this investigation was to compare methods of retrieval for identifying trial protocols and establish:

1. the number of trials that include the trial registration number (TRN) to assess the level of linkage between trial results publication and original protocol.
2. the comprehensiveness of trial protocol sources to identify what proportion of protocols are potentially retrievable.
3. the search methods that are most effective in identifying protocols when registration details are not provided in the trial article to identify the most effective methods for retrieving available protocols.
4. the patterns of protocol registration related to year of publication, funding type, journal type and journal impact factor of the published trial in order to determine

compliance with the International Committee of Medical Journal Editors (ICMJE) requirement for compulsory trial registration in 2005 (De Angelis et al. 2004) and whether any of these factors increase the likelihood that a protocol is registered.

2. Methods

The investigation was conducted over a 4-month period (20th September 2016 to 22nd January 2017). Protocols were searched for 74 randomised controlled trials identified as potential included studies in a systematic review on primary care interventions for medically unexplained symptoms.

2.1 Search Strategy

The search strategy was comprised of 6 stages listed in Table 1. All stages were completed for each protocol for the purpose of this investigation. Each stage had a defined search strategy outlined below. For the purpose of this investigation, trial registry number (TRN) is used as a definition for any unique identification number assigned to a trial, including National Clinical Trial (NCT) number as used in Clinicaltrials.gov.

Stage 1: Protocol Registration Details

Published reports of randomised controlled trials were skim-read to identify protocol registration details. In addition, the terms (1) “regist” (for register/registry/registration), (2) “identif” (for identification/identifier) and (3) “protocol” were searched within the document using the search box as a supplementary check.

Stage 2: Journal Websites

The journal website that published the original trial report was searched using (1) the full trial title as recorded in the trial publication, (2) the full trial title plus the term “protocol” and (3) an abbreviation of the trial title. The first 20 search results were checked for pragmatic reasons, based on relevance ranking.

Stage 3: Internet Searching

The Internet search engine Google was used to search for (1) the full trial title as recorded in the trial publication (2) the full trial title plus the term “protocol”. Again, the first 20 search results were checked for pragmatic reasons, based on relevance ranking.

Stage 4: Trial Registers

The trial registers; Clinicaltrials.gov, International Clinical Trials Registry Platform (ICTRP), International Standard Randomised Controlled Trials Number registry (ISRCTN) and, where applicable, the trial's country-specific register were searched. Where provided in the trial article or if found via any other search method, the TRN was introduced in the search box to confirm registration. Where no TRN was available, the following search approach was used: (1) full trial title as recorded in the trial publication (2) abbreviation of the trial title as recorded in the trial publication (3) simple search for condition combined with the intervention, using the Boolean operator "AND", with supplementary searches when more than one condition and/or intervention were stated on the title (4) advanced search entering the condition in the box labelled "Condition" and the intervention in the box labelled "Intervention", and when results were >20 results the trial's country was also introduced in the "Geographical location" box (Clinicaltrials.gov; ICTRP) or the "Countries of recruitment" box (ISRCTN). In all strategies the first 20 search results were checked for pragmatic reasons, based on relevance ranking.

Stage 5: Bibliographic Databases

The bibliographic databases CENTRAL (via The Cochrane Library) and Ovid MEDLINE were searched using multiple strategies. In CENTRAL (1) the full trial title as recorded in the trial publication (2) an abbreviation of the trial title, (3) the condition combined with the intervention, using the Boolean operator "AND" (with multiple search strategies where more than one condition and/or intervention were stated on the title), were searched in the "Title, Abstract, Keywords" fields, (4) the first author's name was entered in the "Author" box. All punctuation was removed from search terminology in acknowledgement of the sensitivity of the CENTRAL search function to punctuation marks. In Ovid MEDLINE, the advanced "title" search facility was used for the full trial title or an abbreviation of the trial title (both as recorded in the trial publication), and an advanced "multi-field" search was used to enter the condition (selecting "all fields") and the intervention (selecting "all fields"). When more than 20 results were retrieved, this was combined with the primary author's name in the "author" field.

Stage 6: Contacting Authors

Corresponding authors were contacted as part of the review process via email using a standard template developed by the primary investigator (PI) of the systematic review (JL) in consultation with the project team. The email fully cited the index study of interest article and requested the most up-to-date study protocol, allowing an opportunity for authors to forward any updated unpublished versions. This email also requested links to related papers that may not have been identified in the searches but that may have met the inclusion criteria. .

Where contact details were no longer current (e.g. emails bounced back), the Internet was searched for alternative contact details. Responses were recorded over a one-month period. Where respondents attached a copy of, link to, reference to or location for the study protocol these details were considered a positive outcome of the protocol search.

In stages 2-4, differences between UK and US word spelling were considered; terms were always introduced in the search as derived from the trial title. Where author name included hyphen, accent and/or punctuation, the search was performed twice, once entering the name as spelt and once without these. A strict interpretation was used for terms for interventions and conditions from the title, except where clear inferences could be made; for instance, if the condition was “low back pain” searches were also made using “back pain”.

Identification of the correct protocol was verified (1) from the protocol registries, if the published trial was cited in the publications section and (2) by comparing the inclusion and exclusion criteria, outcome measures, sample size and funding source between the identified protocol and the published trial.

2.2 Data Analysis

A Microsoft EXCEL spreadsheet was used to record whether or not the protocol had been retrieved for each individual search strategy, in order to determine the efficacy of each retrieval method and coverage of each source. The funding source, as stated in the published articles, and the journal’s Impact Factor, based on Web of Science’s InCites Journal Citation Reports, were also recorded. The search protocol was developed by one author (AS) in collaboration with the project team. One author (MCG) performed the protocol search and recorded the findings, funding sources and journal Impact Factor. One author (MCG) contacted the authors with contact details of the Principal Investigator of the systematic review (JL) being supplied in case of any queries.

3. Results

3.1 Study Characteristics

74 randomised controlled trials were identified as potentially includable studies for the systematic review at the time of the investigation. The majority of these trials (76% n=56) were published post-2005 (see Figure 1), after the ICMJE compulsory trial registration statement (De Angelis et al. 2004) came into effect.

3.2 Protocol Identification

57% (n=42) of trial protocols were identified following completion of all six stages of searching (see Table 2). At Stage 1 (checking studies for protocol registration details), 43 studies reported the existence of a protocol. 13 protocols were referenced in the study journal article including a full citation for the published protocol, 8 included a trial registration number (TRN), typically located following the journal abstract, but also found in methods sections, acknowledgements and endnotes. 22 referred to a protocol, with varying detail to aid identification of a published protocol. Some studies simply mentioned the existence of a protocol, in conjunction with approval by institutional review boards and ethics committees, without giving further details of separate publication. One study specifically stated that the protocol was available from the authors upon request. Successful identification at Stage 1 was defined for studies that either a) included a full reference to a trial protocol and/or b) included a TRN. Therefore 15 protocols were identified by checking the studies for registration and publication details, all published post-2005. Only one protocol was still found to be uniquely identified via this method following completion of all six stages of searching.

At Stage 2 (searching journal websites) two protocols were identified. Both protocols were distinctively retrieved by an abbreviated trial acronym. Searching by the full trial title, with or without the term “protocol” did not retrieve these references. Both protocols had been identified at Stage 1, therefore no unique protocol references were identified via journal websites. The journals that published the protocols, in addition to the original trial studies, were both BMC titles (Gastroenterology, Psychiatry), neither being ICMJE member journals.

Stage 3 (Internet Searching) identified 23 trial protocols. Eight of the protocols identified via Google searching had not been identified in Stages 1 & 2, however once all six stages of searching had been completed Google searching did not identify any unique protocols. All protocols identified via Google searching were for trials published post-2005. Searching by the title of the published trial was the most effective method, when examining the first 20 results, with 22of the protocols identified via Google retrieved by this method. Searching by the trial title plus the term “protocol” identified 14 of the protocols, with only one uniquely identified by this method.

Stage 4 (Trial Registers) identified 38 protocols across the three sources (clinicaltrials.gov, ICTRP, ISRCTN). 6 of the protocols identified in trial registers were uniquely identified by one source (ICTRP in all cases). 6 of the total number of protocols searched for had not previously been identified in Stages 1-3, and overall 5 were uniquely retrieved via trials register searching. All 38 of the trials identified by the registers were published post-2005. Two of the protocols pre-dated all three registers so would not be available via these sources.

18 of the protocols identified from the trial registers were found in clinicaltrials.gov. All 18 of these protocols were retrieved by TRN. 5 of these were uniquely retrieved from clinical trials.gov using this method. Searching by the trial study title did not retrieve any protocols. Searching variously by "Condition" AND "Intervention" via Basic Search, and by searching for "Condition" AND "Intervention" AND "Geographic Location" in Advanced Search both retrieved 8 of the 18 protocols found via clinicaltrials.gov, but not the same 8 protocols. Searching for "Condition" AND "Intervention" via Advanced Search found 7 of the 18 protocols found. Searching by abbreviation of the trial study title (e.g. +DWR for the "Plus Deep Water Running on Low Back Pain" study) identified 3 of the 18 protocols found.

All 38 of the protocols identified via trial registers were retrieved from ICTRP. 37 of these were identified by searching for the TRN. The remaining protocol was retrieved uniquely by searching for "Condition" AND "Intervention" in the basic search function on the registry search portal homepage, with this method retrieving 15 of the trials found via this source overall. Searching by title, abbreviation, and in Advanced Search for "Condition" AND "Intervention" (with and without geographical location) were much less fruitful. All these methods retrieved between two or three protocols, with no method uniquely identifying any protocols.

13 of the protocols identified via trial registers were retrieved from ISRCTN. All 13 protocols were retrieved by the TRN, although none were identified uniquely by this method. Searching by the trial name abbreviation or "Condition" AND "Intervention" via the search box on the homepage, were comparably effective methods, identifying 11 and 10 of the trials identified via ISRCTN respectively. Searching for "Condition" AND "Intervention" AND "Countries of recruitment" via the "Advanced Search" function identified 7 of the protocols found via this source. Searching for "Condition" AND "Intervention" via "Advanced Search" or searching for the trial study title on the homepage basic search function both identified 6 of the trials found on ISRCTN, but not the same 6.

Country-specific registers offered potential coverage of 5 protocols, and 2 were identified, but neither uniquely. Both these protocols were identified by searching for the TRN or the "Condition" AND "Intervention". Neither were identified by searching for the trial title or abbreviation. For two protocols where a country-specific register was available, it was not possible to search in the English language.

Stage 5 (Database Searching) identified 9 of the protocols searched for, across two sources (CENTRAL and MEDLINE). All nine had already been identified in Stages 1-4. 3 of the protocols found via database searching were uniquely identified via CENTRAL, and 2 were uniquely identified via MEDLINE. All nine protocols identified by database searching related to trials published post-2005.

CENTRAL identified 7 of the protocols retrieved via database searching. The most effective retrieval method utilised the trial name abbreviation in the "Title, Abstract, Keywords" fields, identifying 5 of the protocols identified via CENTRAL. 3 of these were uniquely identified on CENTRAL via this method. Searching by "Condition" AND "Intervention" identified 4 of the protocols retrieved from CENTRAL, with one of these being uniquely identified by this method. Searching by author identified two of the protocols in CENTRAL, but not uniquely. Searching by title did not identify any of the protocols.

MEDLINE identified 6 of the protocols retrieved via database searching. The most effective retrieval method involved searching for the "Condition" AND "Intervention" (in all fields .af) AND "Primary Author Name" in the author field (.au) which identified all six protocols retrieved via MEDLINE, with 4 being identified uniquely via this method. Searching by "Condition" AND "Intervention" alone in all fields, identified two of the protocols in MEDLINE. Searching by the full trial title (in the title field .ti) or searching by the abbreviated title (in .ti) both identified the same single protocol.

In Stage 6 (contacting authors), 18 protocols were identified. 22 authors replied in total. 16 of the email addresses of authors contacted were identified as no longer current, due to emails "bouncing back" with no forwarding email or alternative contact given. 9 of the protocols retrieved via authors were uniquely identified at this stage, with 3 of these being published pre-2005. Of the protocols identified via contact with authors, 11 provided an attached copy of the protocol, 5 directed us to accessing the protocol, either via a direct link or referring to the original trial publication containing the details. Two authors additionally provided the trial registration number, however in both cases this was recorded in the original trial publication (both published post-2005) so had already been identified at this stage. Two of the protocols identified at this stage were published in a language other than English.

Of the 39 protocols officially published and/or registered (relating to trials published post-2005), the majority (n=35) were published prior to publication of the trial results. The recommendation from the ICMJE is that trials published in their member journals are registered prior to patient enrolment, with effect from 1 July 2005 (De Angelis et al. 2004). The remaining 4 protocols were registered/published after publication of trial results. These protocols relate to studies published between 2009-2012. However, only one protocol was published in an ICMJE journal with patient enrollment taking place post-July 2005.

The majority (n=17) of published/registered protocols were government funded. This is consistent with the total number of government-funded trials in this investigation (n=32), therefore may not reflect funding requirements. Journals in which trials were published were

classified as a general medical journal (for example BMJ or PLOS One) or a specialised journal (for example Journal of Psychosomatic Research or Behaviour Research and Therapy). 47 of the trials were published in speciality journals and 27 were published in general medical journals. This is reflected in the proportion of those protocols published and/or registered, with 17 published in general medical journals and 22 published in specialty journals. The Journal Impact Factor (JIF) ranged from 1.061-19.967. One trial was not included in this analysis as the journal did not report the JIF, instead the SCImago Journal Rank was reported (0.507) which is not directly comparable. Similarly, the trials with published/registered protocols were reported in journals with Impact Factors ranging from 1.217-19.967 so a higher Impact Factor was not associated with the registration of trial protocols. 24 of the protocols searched for were related to trials published in ICMJE member journals. Of these 24 protocols, 14 were retrieved, all published post-2005. Of the 10 trials published in ICMJE journals that were not retrieved, 8 were published pre-2005. However, the remaining 2 were published post-2005, in 2009 and 2013 respectively, therefore compliance remains incomplete at the time of this investigation.

43% (n=32) of the protocols remained unidentified after all 6 stages of searching were completed. Half of these (n=16) were for trials published pre-2005.

4. Discussion

Despite conducting searching via 6 sources and using multiple retrieval methods, total coverage was not achieved. Protocols remaining unidentified were split evenly between trials published pre- and post-2005, so this issue is not specific to older trials. Our investigation found that contact via authors was the only effective method of identifying protocols pre-2005. However, given the increased likelihood of contact details no longer being current for older publications, this method remained unsatisfactory when identifying pre-2005 protocols. For example, the earliest included trial publication (published in 1995) stated that the protocol was available from the authors upon request, however the corresponding author did not respond to our request. Searching journal websites was not an effective method for identifying protocols due to limited publication of protocols in the journals searched at the time of the investigation, only two protocols were published in the same journal as the trial and neither were identified uniquely via this method. Database searching had limited effectiveness and was not a unique source of identification for any of the protocols. Checking the trial publication was more effective at identifying the existence of a protocol than in actually facilitating retrieval of the protocol, due to the limited detail within many publications, particularly pre-2005. Checking the trial publication was one of the quickest methods of retrieval in terms of time spent but a relatively low yield, only 13 protocols were cited in the associated trial allowing a direct link between the two publications. Contacting authors was also a relatively quick and straightforward process, using a standard email template for the initial contact. Conversely, searching trial registers

and databases was more time consuming, particularly as multiple search strategies may be required for each trial protocol.

Of the trial registers searched, ICTRP was the most comprehensive source, followed by clinicaltrials.gov, then ISRCTN. Searching country-specific registries where available was not an effective method of retrieval, and revealed issues relating to registers published in languages other than English, therefore searching these sources was not always possible. It is noted that searching additional registry sources not covered by ICRTP may be more useful to obtain a global view (Pansieri et al. 2017) when searching for potential RCTs to be included in systematic reviews rather than searching for known items (trial protocols) as we were in this study.

Internet Searching delivered a moderate yield, but is not effective at uniquely identifying protocols over other sources. Searching databases was not considered an effective method due to a relatively low yield and non-identification of unique protocols. Database searching is also time-consuming due to inadequate coverage of single sources and the multiple search strategies required to retrieve all the protocols indexed by a single source.

When searching trial registers, the trial registry number is the most effective method of retrieval. If the TRN has not been identified, searching by condition and intervention is the next best method. In Internet Searching, the trial publication title is the most effective method of retrieval. It is not possible to recommend a single method of retrieval to identify the majority of protocols when database searching, but searching by the trial name abbreviation was the most effective method of searching CENTRAL. Relatively speaking, searching by condition and intervention and author was an effective method of searching MEDLINE, but overall this database had a relatively low yield from the search methods used. Suboptimal retrieval may occur where the original title of the protocol differs from the final title of the trial report. The most efficient retrieval method for each source can be found in Table 3.

No association was found between the journal type, funding source, and Journal Impact Factor, and the publication and/or registration of trial protocol.

Our study findings correspond to the Glanville et al. (2014) study which explored searching trial registers to inform systematic reviews. Glanville et al. (2014) found a sensitive search approach used in the basic search function of trial registers was the most effective retrieval method, using condition and intervention terms. They reported a relatively low yield from trial registers (on average 16% per systematic review) but note that poor performance could be partly attributed to trials being published pre-2005 (28%). Despite a similar percentage of pre-2005 trials searched for in our study (24%), our yield from the registers was considerably

higher (51%). However, the differences between the two studies must be noted, Glanville et al. (2014) included trials from 8 systematic reviews, all on drug interventions or procedures, compared with our single systematic review on behavioural interventions, which may attribute to the higher yield along with allowance for more trials being added to the registers over time.

Van de Wetering et al. (2012) found that 55-60% of reports of RCTs (retrieved from MEDLINE or the Netherlands Trial Register) contained a trial registration number. This corresponds with our finding that 55% of studies recorded the existence of a protocol, however only 11% contained a TRN. However it is noted that our investigation included pre-2005 reports of trials, whereas Van de Wetering et al. (2012) did not.

Various efforts have been made to improve prospective registration of clinical trials, including the AllTrials initiative (AllTrials 2014) and the requirement to register trials prior to publication in ICJME member journals (De Angelis et al. 2004). Barriers in protocol registration remain for some researchers, including protection of efforts and ideas and decreased autonomy (Moher et al. 2016). This final barrier can be addressed by documenting and reporting changes, for example all primary registries included in the ICTRP are required to provide an audit trail of any changes to trial profiles (Huic et al. 2011). A direct link between clinical trial registrations and their published results is essential to improve efficient identification of trial protocols, and is being improved by the Linked Clinical Trials project which aims to connect all articles relating to an individual trial by its TRN (Shanahan & Meddings 2016). However, despite the growth in registration of trials, bi-directional links between trial registry entry and published results has not increased over time according to a recent systematic review (Bashir et al. 2017).

4.1 Limitations

This investigation was conducted for a single systematic review of behavioural interventions. The complex nature of the topic, the inclusion of trials pre- and post-2005, and consistencies with the existing literature, lead us to cautiously recommend that our findings may be relevant across other reviews, but note that additional sources may be required for systematic reviews of drug interventions, for example pharmaceutical manufacturers' trial databases. Our study is opportunistic and based on a case study approach. Therefore similar studies in different topic areas would aid the assessment of transferability of the findings.

In contacting authors, only the corresponding author was contacted, using the contact details given on the trial publication, unless the corresponding author gave an alternative contact. Where email addresses were no longer valid we sought alternative contact details by Internet searching, but we made no attempt to contact other listed authors on the trial

publications. This approach would be time-consuming with little expectation of response from non-corresponding authors and the ethics of contacting those who have not nominated themselves as available for correspondence might be considered questionable. In addition, six of the authors contacted were responsible for more than one trial, therefore although counted as a single response or non-response per trial, it is expected that an author replying for one trial may respond regarding all their trials. Conversely a non-response might be expected to extend to all trials. However, where contributors to multiple trials responded they did not necessarily provide a protocol for all their trials; the date since publication was important as for other retrieval methods. In addition, one corresponding author was a clinical expert on the review team, with prior knowledge of the review, and was anticipated to be more likely to respond in a timely manner. Due to the time limitations of this investigation, we were unable to subsequently contact non-responding authors as a reminder, this may be an effective approach to identify further protocols, where time and resources allow.

5. Conclusions

Based on our investigation, the following recommendations can be made. Given the limited pre-2005 coverage of sources such as trials registers, to identify the protocols of trials published pre-2005, a review team should contact authors direct as a priority method of retrieval over other search methods. No single source was effective in identifying all protocols post-2005, but prioritisation of retrieval methods can be recommended. This investigation found that searching trial registers was most effective if the TRN has already been identified, and we found that authors contacted were likely to refer to the protocol publication on the trial register if they were contacted prior to checking this source. Therefore a review team should check trial publications for reference to the protocol as the first stage of retrieval. Next, review teams should search trial registers, in the following order until a predetermined point of diminishing returns has been reached; (1) ICTRP (2) clinicaltrials.gov (3) ISRCTN. If the trial registry number has not been identified, the team should conduct a basic search of trial registries by condition and intervention (noting that multiple searches utilising permutations of synonyms may be required). For trials where the protocol has not been identified after these two procedures, a review team should contact corresponding authors via email. If responses from authors is suboptimal, Internet searching, followed by database searching, can be utilised for any remaining unidentified protocols. However these stages are only indicated for trials post-2005, and only if time and resources allow, given that these methods are likely to have a moderate-low yield and are time-intensive to search effectively. Identification of published protocols has been improved by the publication and greater adoption of reporting guidelines such as CONSORT (Schulz et. al 2010), but availability and retrieval remain suboptimal.

Acknowledgements

This study was carried out as part of a project funded by the National Institute of Health Research (NIHR) Health Technology Assessment (HTA) Programme, project number 14/26/08. The views and opinions expressed therein are those of the authors and do not necessarily reflect those of the Health Technology Assessment (HTA) Programme NIHR, NHS or the Department of Health.

The authors thank Sheldon Korpet, Information Officer at School of Health and Related Research (ScHARR), The University of Sheffield, for assistance with reference management.

References:

AllTrials Web site. http://www.alltrials.net/. Published 2014. Updated July 27, 2017. Accessed September 11, 2017.

Abaid, L.N., Grimes, D.A. and Schulz, K.F., 2007. Reducing publication bias through trial registration. *Obstetrics & Gynecology*, *109*(6), pp.1434-1437.

Al-Marzouki, S., Roberts, I., Evans, S. and Marshall, T., 2008. Selective reporting in clinical trials: analysis of trial protocols accepted by The Lancet. *The Lancet*, *372*(9634), pp.201.

Bashir, R., Bourgeois, F. T., and Dunn, A. G., 2017. A systematic review of the processes used to link clinical trial registrations to their published results. *Systematic Reviews*, 6(1), pp.123.

Bashir, R. and Dunn, A.G., 2016. Systematic review protocol assessing the processes for linking clinical trial registries and their published results. *BMJ Open*, *6*(10), pp.e013048.

Baudard, M., Yavchitz, A., Ravaud, P., Perrodeau, E. and Boutron, I., 2017. Impact of searching clinical trial registries in systematic reviews of pharmaceutical treatments: methodological systematic review and reanalysis of meta-analyses. *BMJ*, *356*, pp.j448.

Bourgeois, F.T., Murthy, S. and Mandl, K.D., 2010. Outcome reporting among drug trials registered in ClinicalTrials. gov. *Annals of Internal Medicine*, *153*(3), pp.158-166.

Chalmers, I. and Glasziou, P., 2009. Avoidable waste in the production and reporting of research evidence. *Obstetrics & Gynecology*, *114*(6), pp.1341-1345.

Chan, A.W., 2012. Out of sight but not out of mind: how to search for unpublished clinical trial evidence. *BMJ*, *344*, pp.d8013.

Chan, A.W. and Altman, D.G., 2005. Identifying outcome reporting bias in randomised trials on PubMed: review of publications and survey of authors. *BMJ*, *330*(7494), pp.753.

Chan, A.W., Hróbjartsson, A., Haahr, M.T., Gøtzsche, P.C. and Altman, D.G., 2004a. Empirical evidence for selective reporting of outcomes in randomized trials: comparison of protocols to published articles. *JAMA*, *291*(20), pp.2457-2465.

Chan, A.W., Krleža-Jerić, K., Schmid, I. and Altman, D.G., 2004b. Outcome reporting bias in randomized trials funded by the Canadian Institutes of Health Research. *Canadian Medical Association Journal*, *171*(7), pp.735-740.

Chan, A.W., Hróbjartsson, A., Jørgensen, K.J., Gøtzsche, P.C. and Altman, D.G., 2008. Discrepancies in sample size calculations and data analyses reported in randomised trials: comparison of publications with protocols. *BMJ*, *337*, pp.a2299.

Chan, A.W., Tetzlaff, J.M., Altman, D.G., Dickersin, K. and Moher, D., 2013a. SPIRIT 2013: new guidance for content of clinical trial protocols. *The Lancet*, *381*(9861), pp.91.

Chan, A.W., Tetzlaff, J.M., Gøtzsche, P.C., Altman, D.G., Mann, H., Berlin, J.A., Dickersin, K., Hróbjartsson, A., Schulz, K.F., Parulekar, W.R. and Krleža-Jerić, K., 2013b. SPIRIT 2013 explanation and elaboration: guidance for protocols of clinical trials. *BMJ*, *346*, pp.e7586.

De Angelis, C., Drazen, J.M., Frizelle, F.A., Haug, C., Hoey, J., Horton, R., Kotzin, S., Laine, C., Marusic, A., Overbeke, A.J.P. and Schroeder, T.V., 2004. Clinical trial registration: a statement from the International Committee of Medical Journal Editors. *New England Journal of Medicine*, *351*(12), pp.1250-1251.

Dwan, K., Altman, D.G., Arnaiz, J.A., Bloom, J., Chan, A.W., Cronin, E., Decullier, E., Easterbrook, P.J., Von Elm, E., Gamble, C. and Ghersi, D., 2008. Systematic review of the empirical evidence of study publication bias and outcome reporting bias. *PLOS ONE*, *3*(8), pp.e3081.

Dwan, K., Altman, D.G., Cresswell, L., Blundell, M., Gamble, C.L. and Williamson, P.R., 2011. Comparison of protocols and registry entries to published reports for randomised controlled trials. *Cochrane Database of Systematic Reviews*, Issue 1. Art. No.: MR000031. DOI: 10.1002/14651858.MR000031.pub2.

Fredrickson, M.J. and Ilfeld, B.M., 2011. Prospective trial registration for clinical research: what is it, what is it good for, and why do I care?. *Regional Anesthesia and Pain Medicine*, *36*(6), pp.619-624.

Getz, K.A., Zuckerman, R., Cropp, A.B., Hindle, A.L., Krauss, R. and Kaitin, K.I., 2011. Measuring the incidence, causes, and repercussions of protocol amendments. *Drug Information Journal*, *45*(3), pp.265-275.

Gill, C.J., 2012. How often do US-based human subjects research studies register on time, and how often do they post their results? A statistical analysis of the Clinicaltrials. gov database. *BMJ Open*, *2*(4), pp.e001186.

Glanville, J.M., Duffy, S., McCool, R. and Varley, D., 2014. Searching ClinicalTrials. gov and the International Clinical Trials Registry Platform to inform systematic reviews: what are the optimal search approaches?. *Journal of the Medical Library Association*, *102*(3), pp.177-183.

Glasziou, P., Altman, D.G., Bossuyt, P., Boutron, I., Clarke, M., Julious, S., Michie, S., Moher, D. and Wager, E., 2014. Reducing waste from incomplete or unusable reports of biomedical research. *The Lancet*, *383*(9913), pp.267-276.

Hartung, D.M., Zarin, D.A., Guise, J.M., McDonagh, M., Paynter, R. and Helfand, M., 2014. Reporting Discrepancies Between the ClinicalTrials. gov Results Database and Peer-Reviewed Publications. *Annals of Internal Medicine*, *160*(7), pp.477-483.

Huić, M., Marušić, M. and Marušić, A., 2011. Completeness and changes in registered data and reporting bias of randomized controlled trials in ICMJE journals after trial registration policy. *PLOS ONE*, *6*(9), pp.e25258.
Huynh, L., 2011. Primary outcomes reported in abstracts and ClinicalTrials.gov - do they agree? Oral presentation at the 19th Cochrane Colloquium; Oct 19-22; Madrid, Spain. *Cochrane Database of Systematic Reviews*, Supplement 18.
Ko, H., Tai, F.M., Ghersi, D. and Askie, L.M., 2011. Inconsistent quality of reporting of searching clinical trials registries in Cochrane systematic reviews and protocols. Poster presentation at the 19th Cochrane Colloquium; Oct 19-22; Madrid, Spain. *Cochrane Database of Systematic Reviews*, Supplement 56.
Law, M.R., Kawasumi, Y. and Morgan, S.G., 2011. Despite law, fewer than one in eight completed studies of drugs and biologics are reported on time on ClinicalTrials. gov. *Health Affairs*, *30*(12), pp.2338-2345.
Lefebvre, C., Glanville, J., Wieland, L.S., Coles, B. and Weightman, A.L., 2013. Methodological developments in searching for studies for systematic reviews: past, present and future?. *Systematic Reviews*, *2*(1), pp.78.
Moher, D., Glasziou, P., Chalmers, I., Nasser, M., Bossuyt, P. M., Korevaar, D. A., Graham, I.D., Ravaud, P. and Boutron, I., 2016. Increasing value and reducing waste in biomedical research: who's listening?. *The Lancet*, 387(10027), pp.1573-1586.
Nurbhai, M., Grimshaw, J., Lorenzo, M.P., Liberati, A., Chan, A., Dickersin, K., Krezla-Jeric, K. and Moher, D., 2005. *Assessing the quality of information recorded on trial registries*. Cochrane Colloquium; Oct 22-26; Melbourne, Australia.
Odutayo, A., Altman, D.G., Hopewell, S., Shakir, M., Hsiao, A.J. and Emdin, C.A., 2015. Reporting of a publicly accessible protocol and its association with positive study findings in cardiovascular trials (from the Epidemiological Study of Randomized Trials [ESORT]). *The American Journal of Cardiology*, *116*(8), pp.1280-1283.
Pandis, N., Fleming, P.S., Koletsi, D. and Hopewell, S., 2016. The citation of relevant systematic reviews and randomised trials in published reports of trial protocols. *Trials*, *17*(1), pp.581.
Pansieri, C., Pandolfini, C., and Bonati, M., 2017. Clinical trial registries: more international, converging efforts are needed. *Trials*, 18(1), pp.86.
Scherer, R.W., Sieving, P.C., Ervin, A.M. and Dickersin, K., 2012. Can we depend on investigators to identify and register randomized controlled trials?. *PLOS ONE*, *7*(9), pp.e44183.
Schroll, J.B., Bero, L. and Gøtzsche, P.C., 2013. Searching for unpublished data for Cochrane reviews: cross sectional study. *BMJ*, *346*, pp.f2231.
Schulz, K.F., Altman, D.G. and Moher, D., 2010. CONSORT 2010 statement: updated guidelines for reporting parallel group randomized trials. *Annals of Internal Medicine*, *152*(11), pp.726-732.

Shanahan, D. and Meddings, K. Clinical trial data and articles linked for the first time. Crossref Blog. https://www.crossref.org/blog/linked-clinical-trials-are-here/. Published May 17, 2016. Accessed September 11, 2017.

Tai, F. M., Willson, M.L., and Ghersi, D. 2012. Implications of searching multiple trial registries: how should we search ClinicalTrials.gov and WHO ICTRP? Cochrane Colloquium. Auckland NZ. Available at: http://2012.colloquium.cochrane.org/abstracts/implications-searching-multiple-trial-registries-how-should-we-search-clinicaltrialsgov-an.html [Accessed 24 February 2017].

Tetzlaff, J.M., Chan, A.W., Kitchen, J., Sampson, M., Tricco, A.C. and Moher, D., 2012. Guidelines for randomized clinical trial protocol content: a systematic review. *Systematic Reviews*, *1*(1), pp.43.

van de Wetering, F.T., Scholten, R.J., Haring, T., Clarke, M. and Hooft, L., 2012. Trial registration numbers are underreported in biomedical publications. *PLOS ONE*, *7*(11), pp.e49599.

Viergever, R.F. and Ghersi, D., 2011. The quality of registration of clinical trials. *PLOS ONE*, *6*(2), pp.e14701.

Viergever, R.F. and Li, K., 2015. Trends in global clinical trial registration: an analysis of numbers of registered clinical trials in different parts of the world from 2004 to 2013. *BMJ Open*, *5*(9), pp.e008932.

Wolfe, N., Gøtzsche, P.C. and Bero, L., 2013. Strategies for obtaining unpublished drug trial data: a qualitative interview study. *Systematic Reviews*, *2*(1), pp.31.

Young, T. and Hopewell, S., 2011. Methods for obtaining unpublished data. *Cochrane Database of Systematic Reviews* 2011, Issue 11. Art. No.: MR000027. DOI: 10.1002/14651858.MR000027.pub2.

Zarin, D.A., Tse, T., Williams, R.J., Califf, R.M. and Ide, N.C., 2011. The ClinicalTrials. gov results database—update and key issues. *New England Journal of Medicine*, 364, pp.852-860.

Table 1: Outline of the 6 stages of protocol identification

Stage 1	Report of protocol registration details in the published articles was checked
Stage 2	The journal websites for the published study were searched for publication of the corresponding protocol
Stage 3	Internet searching
Stage 4	Trial registers
Stage 5	Bibliographic databases
Stage 6	Contacting authors

Table 2: Coverage of Sources

Total protocols identified	**57%**	
By Method		**Unique protocols identified**
Trial Registers	51%	7%
Internet Searching	31%	0%
Contact with Authors	24%	12%
Checking Trial Publication	20%	1%
Database Searching	12%	0%
Journal Websites	3%	0%

Table 3: Efficiency of Search Strategies by Source

Source	Search Method	Retrieval
Internet Searching (Google)	Title	96%
Trial Registers (ICRTP, clinicaltrials,gov, ISRCTN)	Trial Registry ID	97-100%
Database Searching (CENTRAL)	Trial Name Abbreviation	71%
Database Searching (MEDLINE)	Condition AND Intervention AND Author	100%

Figure 1: Year of Trial Publication