Transparency of outcome reporting and trial registration of randomized controlled trials in top psychosomatic and behavioral health journals: A 5-year follow-up

Running Head: Outcome Reporting and Trial Registration

Kira E Riehm[1,2]; Marleine Azar[1,2]; Brett D Thombs, PhD[1-7]

[1]Lady Davis Institute for Medical Research, Jewish General Hospital, Montréal, Québec, Canada; [2]Departments of Psychology, [3]Psychiatry, [4]Epidemiology, Biostatistics, and Occupational Health, [5]Educational and Counselling Psychology, [6]Medicine,and [7]School of Nursing, McGill University, Montréal, Québec, Canada.

Address for Correspondence: Brett D Thombs, PhD; Jewish General Hospital; 4333 Cote Ste Catherine Road; Montreal, Quebec H3T 1E4; Tel: (514) 340-8222 ext. 5112; Email: brett.thombs@mcgill.ca

Financial support: Dr. Thombs is supported by an Investigator Salary Award from the Arthritis Society. There was no direct financial support for the study, and there are no conflicts of interest to report.

Abstract

Objective: The extent that randomized controlled trials (RCTs) accurately reflect intervention effectiveness depends on the completeness and accuracy of published results. A previous study found that only 40% of 63 RCTs published in top behavioral health journals in 2008-2009 clearly declared primary and secondary outcomes and only 21% were registered. The objective of this study was to conduct a five-year follow-up to assess outcome reporting clarity, proportion of registered trials, and adequacy of outcome registration in RCTs in top behavioral health journals.

Method: Eligible studies were RCTs published in *Annals of Behavioral Medicin*e, *Health Psychology*, *Journal of Psychosomatic Research*, and *Psychosomatic Medicine* from January 2013 to October 2014.

Results: Of 76 RCT publications reviewed, only 25 (32.9%) adequately declared primary or secondary outcomes, whereas 51 (67.1%) had multiple primary outcomes or did not define outcomes. Of the 76 trials, 40 (52.6%) had been registered. Only 3 studies registered a single primary outcome and time point of assessment prior to enrolling patients, and registered and published outcomes were discrepant in 1 of the 3 studies. No studies were adequately registered as per Standard Protocol Items: Recommendation for Interventional Trials guidelines. Compared to 5 years prior, the proportion of published trials with adequate outcome declaration decreased from 39.7% to 32.9% ($p = 0.514$). The proportion of registered trials increased from 20.6% to 52.6% ($p < 0.001$).

Conclusion: The quality of published outcome declarations and trial registrations remains largely inadequate. Greater attention to trial registration and outcome definition in published reports is needed.

Keywords: behavioral medicine; bias; randomized controlled trials; selective outcome reporting; trial registration

INTRODUCTION

Well-designed and conducted randomized controlled trials (RCTs) provide the highest quality evidence for evaluating the effectiveness of health care interventions [1]. The degree to which published reports of RCTs reflect accurate, realistic estimates of intervention effectiveness, however, depends on how outcomes are defined and reported [2-4].

Most RCTs assess multiple outcomes, but ideally a single primary outcome variable is identified to answer the main question the trial is designed to address. Other outcome variables are designated as secondary [3-5]. The failure to designate a primary outcome complicates interpretation of results, particularly when different outcome variables give contradictory results [3]. Furthermore, without statistical adjustment, multiple outcomes generate a potentially large number of hypothesis tests, which increases the likelihood of false-positive claims of effectiveness [3, 6].

Not designating primary outcomes and analysis methods *a priori* can also lead to selective outcome reporting [3, 7], which occurs when statistically significant or outcomes from a study are published, whereas negative outcomes from the same study are not [8]. Even when a single primary outcome variable is identified, selective reporting biases may occur if multiple analyses methods are undertaken, if outcomes are assessed at multiple time points, or if outcome variables are compared using different metrics (e.g., change from baseline, final value) or methods of aggregation (e.g., mean, median) without specification of the primary method prior to enrolling patients [9, 10]. The existence of selective reporting biases in published trial reports is well-documented. One review evaluated 16 studies that compared trial protocols or trial registry entries to published results in a median of 54 RCTs (range 2 to 362) and found that at

least one primary outcome had been changed, introduced, or omitted post hoc in 2% to 50% of the trials examined in the 16 studies [8].

Two initiatives have been introduced to increase transparency in trial reporting. First, the Consolidated Standards of Reporting Trials (CONSORT) statement [11] was developed to improve reporting of trials and to guide readers, peer reviewers, and editors in critical evaluations of RCT reports. CONSORT provides authors with a checklist of critical items that should be included in trial reports. A recent review found that RCTs published in medical journals that endorse the CONSORT statement are more completely reported than RCTs published in journals that have not formally endorsed CONSORT [12].

Second, in September 2004, the International Committee of Medical Journal Editors (ICMJE) implemented as policy the requirement that clinical trials must be registered in a public trial registry to be considered for publication in member journals [13]. Ongoing trials that began enrollment prior to July 1st, 2005, were required to register before September 13th, 2005. Trials beginning enrollment after July 1st, 2005, are required to register before beginning patient enrollment. Ongoing trials are defined as trials that are still collecting, cleaning, or analyzing data. Thus, currently, all newly published trials should be registered [13].

Trial registration may reduce publication bias, which occurs when entire trials go unpublished due to unfavorable results, because registration generates a public record of a trial, even if the results are not published [14]. However, trial registration can only reduce selective reporting of some results and not others to the extent that investigators adequately define primary and secondary outcomes in pre-trial registration. ICMJE policy states that, minimally, investigators must define a single primary outcome with a time point of assessment, as well as key secondary outcomes, at the time of registration [15].

In 2011, Milette et al. [16] examined the extent to which RCTs published in 4 top psychosomatic and behavioral medicine journals (*Annals of Behavioral Medicine, Health Psychology, Journal of Psychosomatic Research,* and *Psychosomatic Medicine*) between January 2008 and October 2009 had clearly defined primary and secondary outcomes in the published trial reports and had adequately registered trial outcomes pre-trial. Of 63 published RCTs, only 40% clearly declared primary and secondary outcomes and only 21% were registered. Only 1 trial registered primary outcomes with enough information for comparison to published outcomes, and registered and published outcomes were discrepant. Of the 4 journals reviewed, *Annals of Behavioral Medicine* [17], *Health Psychology* [18], and *Psychosomatic Medicine* [19] began requiring adherence to CONSORT in 2002 or 2003, and the *Journal of Psychosomatic Research* implemented CONSORT following publication of the Milette et al. study in 2011 [20]. Trial registration policies were implemented by *Annals of Behavioral Medicine* in 2010 [21] and by *Psychosomatic Medicine* [22] and the *Journal of Psychosomatic Research* [20] in 2011. *Health Psychology* does not require clinical trial registration.

In 2013, the Standard Protocol Items: Recommendation for Interventional Trials (SPIRIT) statement [23] was published for clinical trial protocols. The SPIRIT guidelines require that for each primary and secondary outcome trial protocols specify: (1) the specific measurement variable (e.g., Beck Depression Inventory score); (2) the participant-level analysis metric (e.g., change from baseline, final value, time to event); (3) the method of aggregation (e.g., mean, proportion above or below a cutoff threshold); and (4) the primary time point of interest for analysis.

The aim of the present study was to provide a five-year update on outcome reporting and trial registration practices for RCTs testing interventions designed to improve health published in

Annals of Behavioral Medicine, Health Psychology, Journal of Psychosomatic Research, and *Psychosomatic Medicine*. Specific objectives were to (1) determine the proportion of RCT publications that clearly defined primary and secondary trial outcomes; (2) assess the proportion of adequately registered RCTs according to the methods used in Milette et al. [16] and the SPIRIT 2013 guidelines; (3) evaluate whether published primary outcomes were consistent with registered primary outcomes; and (4) compare the proportion of published RCTs with clearly defined outcomes and adequate pre-trial registration to results obtained by Milette et al. [16] 5 years ago.

METHODS

Article Selection

Milette et al. [16] reviewed articles published between January 2008 and October 2009 in 4 journals that had been identified previously as "leading psychosomatic and behavioral medicine journals" (p. 206) (*Annals of Behavioural Medicine, Health Psychology, Journal of Psychosomatic Research,* and *Psychosomatic Medicine*) [24]. In the present study, we sought RCTs published in the same 4 journals 5 years later, from January 2013 through October 2014. Titles and abstracts for all articles published in these journals during this period were uploaded into the citation management database RefWorks then into the systematic review manager DistillerSR. DistillerSR was used for all coding procedures and for tracking results of the review process.

Two reviewers independently reviewed titles and abstracts for eligibility. If either reviewer determined that a study was potentially eligible, full-text review was conducted by two reviewers, with disagreement resolved by consensus, including a third investigator as necessary.

Based on the ICMJE definition of clinical trials [13], which has been used in previous studies of RCT registrations [16, 25], articles were included if they reported data from any RCT that randomly assigned participants to intervention and comparison groups to study the cause-and-effect relationship between an intervention and a health outcome. Articles that reported analyses of secondary trial outcomes, including subgroup analyses, were included. Studies that randomized participants into experimental conditions not intended to improve health (e.g., laughter versus mental stress conditions to assess arterial stiffness) or that primarily assessed intervention feasibility were excluded. Articles that reported only mediation or moderation analyses without reporting previously unpublished trial outcomes, used RCT data for cross-sectional analyses only, reported on longitudinal outcomes for all participants in a trial regardless of group assignment, assessed cost-effectiveness, or analyzed only control or treatment group data were excluded.

Data Extraction and Classification

Two investigators independently extracted and entered data into an online database using DistillerSR software. Discrepancies were resolved by consensus.

Objective 1 - clearly and adequately declared outcomes in published articles: Published articles were classified as reporting: (1) *primary*, (2) *multiple primary (same report)*, (3) *multiple primary (different report)*, (4) *multiple primary (with statistical adjustment)*, (5) *secondary*, or (6) *undefined* outcomes.

An article was classified as reporting a *primary* outcome if a single outcome was clearly and consistently defined as primary throughout the article or, alternatively, if a single primary outcome could be determined from the power analysis. Articles that measured a primary outcome variable at multiple time points in the context of a single repeated measures assessment

with only one hypothesis test were classified as reporting a single *primary* outcome. Studies that identified more than one variable as the primary outcome variable or that identified a single variable, but analyzed multiple time points without specification of primacy, were classified as reporting *multiple primary (same report)* outcomes with one exception. If these studies made appropriate statistical adjustments for multiple comparisons, they were classified as reporting *multiple primary (with statistical adjustment)* outcomes. Studies were classified as reporting *multiple primary (different report)* outcomes if they identified a single primary outcome as per the definition of *primary* outcomes, but a previous report had also declared one or more primary outcomes. We identified previous reports by reviewing the references in the included RCT and searching PubMed and PsycInfo using author names and keywords. Studies classified as *primary*, *multiple primary (same report)*, *multiple primary (different report)*, and *multiple primary (with statistical adjustment)* may have also reported secondary outcomes, but this was not recorded.

Studies were classified as reporting *secondary* outcomes if the authors clearly and consistently defined one or more outcomes as secondary and did not report any primary outcomes. Studies were also classified as reporting *secondary* outcomes if there was a clear statement indicating that the primary or main findings of the RCT had been reported in a previous article.

Studies that did not clearly define outcomes as being primary or secondary were classified as reporting *undefined* outcomes. Studies that noted the existence of a previous report, but did not classify outcomes from the previous or current report as primary or secondary, were classified as reporting *undefined* outcomes (e.g., a report of 12-month post-intervention outcomes with a previous report on 6-month outcomes).

Studies with *primary*, *multiple primary (with statistical adjustment)*, or *secondary* outcomes were classified as having *adequately declared* outcomes, whereas studies with *multiple primary (same report)*, *multiple primary (different report)* or *undefined* outcomes were classified as having *inadequately declared* outcomes. We used a Fisher's exact test to determine if the proportion of *adequately declared* outcomes differed between the 4 journals, since there were < 5 expected count per cell in some cases. In the case of statistical significance, post-hoc tests were done to compare two journals at a time with Bonferroni correction for multiple tests (alpha = 0.0083).

Objective 2 - trial registration: We followed a procedure outlined by Mathieu et al. [25], which was also used by Milette et al. [16]. First, we attempted to retrieve trial registration data, including registration number, from each published article. If no registration information was listed in the published article, we contacted the corresponding author by email to determine if the trial had been registered and to obtain the registry name and number, if registered. If no response was received from the corresponding author after 3 contact attempts, each 1 week apart, we searched for the studies in multiple clinical trial registries, including ClinicalTrials.gov (clinicaltrials.gov), ISRCTN (www.isrctn.com), the WHO registry search portal (apps.who.int/trialsearch), and the registry from the country of the first author (e.g., Netherlands Trial Register [www.trialregister.nl]). To identify registry records, we performed a search using key terms from the published article, and then attempted to match the principal investigator, funding source, intervention, control group, and design from the article to the registrations obtained in the search. If this did not uncover a registration number, the published article was coded as not registered. Proportions of registered trials were compared across journals using a Fisher exact test with alpha = 0.05, since expected cell count was < 5 in some cases. In the case

of statistical significance, post-hoc tests were done to compare two journals at a time with Bonferroni correction for multiple tests (alpha = 0.0083).

For registered trials, trial start and end dates and participant enrollment dates were extracted from the publication to determine if the trial was required to have been registered prior to enrolment of any patients per ICMJE policy. If not provided in the publication, start and end dates were extracted from the registration record. For date of registration, for studies in the ISRCTN registry, the "date applied" was extracted. For studies in the ClinicalTrials.gov registry, the "first received" date was extracted. From the registration information, we determined the proportion of published RCTs that were registered versus not registered.

We assessed the adequacy of primary outcome declaration in trial registrations for all published RCTs, except those that only reported secondary outcomes in the publication, using two methods. First, using the same method as Milette et al. [16], registered studies were classified as having *adequately* or *inadequately registered* outcomes. *Adequately registered* studies had to meet 2 criteria, based on ICMJE policy: (1) Studies ongoing as of July 1st, 2005 had to be registered prior to September 13th, 2005 and prior to trial completion. Studies that started after July 1st, 2005 must have been registered before participant enrollment began; (2) Studies had to specify one primary outcome variable in the registration, with a clear description and time frame of assessment, or multiple primary outcomes with a plan for statistical adjustment.

Second, we evaluated the completeness of registered primary outcomes according to SPIRIT 2013 guidelines [23]. For each registered outcome, we evaluated whether the following elements were provided; (1) the specific measurement variable (e.g., Beck Depression Inventory score, not simply "depression"); (2) the participant-level analysis metric (e.g., change from

baseline, final value, time to event); (3) the method of aggregation (e.g., mean, proportion above a cutoff); and (4) the time point for analysis.

With both methods, trial registrations that specified more than 1 primary outcome were considered to have adequately registered the outcome variable only if a plan for statistical adjustment for multiple comparisons was included in the registration or if the same set of primary outcomes were published with proper statistical adjustment. Trial registrations that did not specify a single primary time point for analysis were considered to have met the time point criterion only in the context of a planned analysis that examined change across all time points with a single analysis that tested a single hypothesis. If there were changes in the study registration records, we extracted data from the last registration update prior to initiation of participant enrolment.

Objective 3 - comparison of registered primary outcomes to published primary outcomes: We compared registered primary outcomes to published primary outcomes for each *adequately registered* study to assess consistency of reporting. First, studies classified as *adequately registered* were classified as reporting *consistent* outcomes if the published primary outcomes were identical to the primary outcomes specified in the trial registration. Studies were classified as reporting *discrepant* outcomes if the published primary outcomes differed from the primary outcomes in the trial registration.

Objective 4 - comparison of results to Milette et al. [16]: We compared the proportion of studies with *adequately declared* published outcomes and the proportion of studies that were *adequately registered* in the present study to the results obtained by Milette et al [16]. Proportions of published outcome declarations were compared with a chi-square test, whereas

proportions of adequate registrations were compared using a Fisher exact probability test, given that there were expected cell counts < 5. For both tests, alpha = 0.05.

RESULTS

Article Selection

A total of 955 articles were published in the 4 included journals between January 2013 and October 2014. Of these, 853 articles were excluded at the title/abstract level, leaving 102 articles for full-text review. There were 26 articles excluded after full-text review, leaving 76 articles that reported outcomes from RCTs (Figure 1): 18 from *Annals of Behavioral Medicine*, 33 from *Health Psychology*, 10 from the *Journal of Psychosomatic Research*, and 15 from *Psychosomatic Medicine*. The outcome declaration and registration classifications for all included articles are found in the Appendix.

Objective 1 - clearly and adequately declared outcomes in published articles: Only 25 of 76 articles (32.9%) were classified as having *adequately declared* outcomes, including 17 (22.4%) with *adequately declared primary* outcomes, 2 (2.6%) with *adequately declared multiple primary (with statistical adjustment)* outcomes, and 6 (7.9%) with *adequately declared secondary* outcomes. Of the 51 articles (67.1%) with *inadequately declared* outcomes, 29 (38.2%) declared *multiple primary* outcomes without appropriate statistical adjustment, 19 (25.0%) had *undefined* outcomes, and 3 (3.9%) declared a *primary* outcome, but a previous report from the same RCT declared a different *primary* outcome (Table I). The percentage of *adequately declared* outcomes did not differ significantly across journals (p = 0.798) and ranged from 26.7% in *Psychosomatic Medicine* to 39.4% in *Health Psychology*.

Objective 2 - trial registration: Data on registration status of reviewed RCT articles are shown in Table II. Of the 76 articles reviewed, 40 (52.6%) reported on registered RCTs. Of the

40 articles reporting on a registered RCT, 33 (82.5%) provided registration information in the publication. There was a statistically significant difference in the percentage of registered trials across journals ($p = 0.017$). In post hoc comparisons, the only significant difference was between *Health Psychology* (33.3%) and *Annals of Behavioral Medicine* (77.8%; $p = 0.003$). There were no other statistically significant differences between journals.

As shown in Table II, among the 40 registered trials, 39 were required to be registered pre-enrolment and were considered in our assessment of registration adequacy. Of these, 4 articles reported on only *adequately declared secondary* outcomes. Of the other 35 articles, there were 16 that registered pre-enrolment, but only 3 of the 35 (8.6%) clearly defined a primary outcome and were classified as registering *adequately registered* outcomes per Milette et al. [16] methods. Of the 19 that did not register prior to enrollment, 4 (11.4%) did register a single primary outcome and time point and 15 (42.9%) did not. There were 13 that neither registered pre-enrollment nor registered a primary outcome (37.1%).

Based on the SPIRIT guidelines evaluation of trial registrations [23], there were 7 trials (20.0%) that registered pre-enrolment and adequately specified a primary outcome variable; 3 (8.6%) with pre-enrolment registration, a primary outcome variable, and time point of interest for analysis; 1 (2.9%) with pre-enrolment registration, a primary outcome variable, time point, and analysis metric; but none with all of these plus a method of aggregation.

Objective 3 - comparison of registered primary outcomes to published primary outcomes: Of the 3 articles reporting on RCTs with *adequately registered* outcomes per Milette et al. [16] criteria, 2 articles [26, 27] published outcomes in a manner that was *consistent* with registered outcomes. The remaining article [28] reported outcomes *discrepant* with registered outcomes. The registry entry for this RCT specified the 3-month time point as primary.

However, the publication reported on outcomes for both the 3-month and 6-month time points without specification of primacy.

Objective 4 - comparison of results to Milette et al. [16]: The proportion of published RCTs with *adequately declared* outcomes in the present study (25 of 76; 32.9%) was lower than in the same journals 5 years prior (25 of 63; 39.7%), although the difference was not statistically significant ($p = 0.514$). A significantly greater proportion of trials were registered in the present study (40 of 76; 52.6%) compared to 5 years ago (13 of 63, 20.6%; $p < 0.001$). Only 3 of 76 trials (3.9%) were *adequately registered* in 2013-2014 compared to 1 of 63 (1.6%) in 2008-2009 ($p = 0.626$).

DISCUSSION

Less than a third of RCTs published in 4 top psychosomatic and behavioral health journals in 2013-2014 had adequately declared outcomes in the published reports. Although approximately half of the articles were registered, the quality of registry entries was generally low. Only 3 of 76 published RCTs registered a clearly specified primary outcome variable and time point, and there were no published RCTs that had been registered adequately according to all SPIRIT criteria for trial protocols [23]. Of the 3 articles registered adequately per Milette et al. [16] criteria, 2 reported published outcomes that were consistent with registered outcomes.

Milette et al. [16] found that in a sample of 63 RCTs published in the same 4 journals from January 2008 to October 2009, 40% had adequately declared primary or secondary outcomes in the publication compared to only 33% in the present study. The proportion of published RCTs that had been registered increased from 21% to 53%, an increase of more than 30% from 5 years prior. This is an optimistic finding, which might reflect increased recognition of the need for trial registration by researchers [29], implementation of trial registration policies

in 3 of 4 of the included journals [20-22], or a combination. On the other hand, Milette et al. [16] found only 1 article that had registered sufficient outcome information to compare to published outcomes, and the present study found only 3 such articles. Thus, while the quantity of behavioral health RCTs that are registered has increased in the past 5 years, the quality of trial registrations and outcome reporting has not changed substantively.

In the present study, we did not find studies included in our analyses that were registered prior to patient enrolment and met all SPIRIT criteria. Only 1 trial registration defined an analysis metric and none defined a method of aggregation for clearly specified outcomes. Although research on selective reporting has typically focused on selective reporting of outcome variables, different options available for aggregating results, variable metrics, and analyses can also lead to reporting biases and overly optimistic reports of results from RCTs [10]. Therefore, explicitly outlining any plans for statistical adjustment and analysis in trial protocols is as important as defining a primary outcome and time point for analysis. The SPIRIT guidelines represent an important opportunity for researchers conducting behavioral health trials to improve the quality of protocols and trial registrations [30].

Consistent with the ICMJE definition, we defined an RCT as a comparative study, with random assignment of participants, which tested an intervention intended to improve health outcomes [25]. We included RCTs evaluating interventions that were tested in both clinical and experimental settings, provided that results were reported for an “intervention(s) intended to improve health”. Some might argue that trials done with university student samples and trials that are preliminary tests of intervention mechanisms constitute preclinical research and should therefore be exempt from trial registration requirements. However, ICMJE requirements do not make this distinction. Furthermore, robust evidence shows that the results of a significant

proportion of preclinical studies cannot be replicated [31-35]. This is at least in part due to the susceptibility of preclinical research to many of the same biases that diminish the value of RCTs, generally [36]. Irreproducible studies are a major source of waste in research [31], and the need to improve transparency and reporting in preclinical studies cannot be overemphasized. In fact, the registration of confirmatory preclinical research has previously been put forward as one method of addressing this problem [37-39]. Thus, whether trials of interventions intended to improve health, which are published in top behavioral health journals, are construed as clinical or preclinical trials, it is imperative that they be registered. Greater transparency in the reporting of all research will ultimately increase the likelihood of obtaining reproducible findings.

At present, all journals included in this study either require or request that authors utilize the CONSORT guidelines [4] when reporting results from RCTs. This is an important first step, as research indicates that reports published in journals endorsing this statement are more complete than those published in journals that do not endorse it [12]. However, our results indicate that this initiative alone is not sufficient to ensure the publication of results from trials that appropriately test trial hypotheses via adequate outcome variable definitions and are at low risk of selective outcome reporting. Moving forward, greater attention from authors, peer reviewers, and journal editors is needed to ensure that researchers submitting trial reports for publication are following these guidelines.

With regards to trial registration, *Annals of Behavioral Medicine*, *Journal of Psychosomatic Research*, and *Psychosomatic Medicine* currently have trial registration policies in place. Nonetheless, the proportion of trials published in these journals that are registered remains suboptimal. More active enforcement by peer reviewers and journal editors is needed if the potential of these policies to counteract reporting biases is to be fully realized. Trial

registration alone is not enough to ensure transparent reporting. Only adequate registration can help to ensure full disclosure of trial results in a way that is consistent with the actual conduct of trials. Reviewers and editors should not only ensure that trials are registered, but should also compare trial registrations to submitted trial reports and ask authors to justify any major discrepancies.

Currently, *Health Psychology* is the only of the 4 journals that does not have a trial registration policy in place. Consistent with this, *Health Psychology* had a substantially lower rate of trial registration among published RCTs than the other 3 journals. We recommend that *Health Psychology* adopt and actively enforce a trial registration policy. Additionally, *Health Psychology* and the other 3 journals should attend more closely to the quality of trial registrations. RCTs with poor quality registrations can be identified as at risk of selective outcome reporting, and journals should incorporate this into the judgment of appropriateness and priority for publication.

The burden of ensuring transparent reporting does not rest solely with journal editors or peer reviewers. Despite decades of research that demonstrate vast amounts of waste in research due to publication and selective reporting bias [29, 31, 32], many researchers still appear to be largely unaware of the negative impacts of these biases on the evidence base [40]. Regardless of whether the biased reporting of trial results is due to a poor understanding of trial principles or intentional efforts to mislead, investigators have a fundamental obligation to transparently report the results of RCTs [41]. Using the available reporting guidelines and prospectively registering trials are ways that researchers can fulfill this obligation.

Beyond individual trial reports, it is clear that systematic reviews and meta-analyses based only on published literature may generate inflated effect estimates and draw biased

conclusions [42, 43]. Echoing suggestions made by other authors [44], we recommend that researchers conducting reviews of psychosomatic and behavioral health RCTs include results from both published and unpublished literature. In addition, those who conduct reviews should make use of the Cochrane Collaboration's tool for assessing risk of bias [45], which includes an assessment of selective outcome reporting. In our study, only the 2 articles reporting outcomes consistent with registered outcomes [26, 27] would have been classified as having a low risk of bias; all other articles would have been classified as having either a high or unclear risk of bias. Including unpublished literature and assessing risk of bias for articles included in systematic reviews and meta-analyses will contribute to a more accurate understanding of the effects that are to be obtained from behavioral health interventions.

In the journals that we reviewed for the present study, endorsement of CONSORT and trial registration requirements did not translate into adequate reporting of trial outcomes that could be compared to clearly defined outcomes in trial registrations. A potential solution might be for journals to insist that peer reviewers and editors routinely review the adequacy of outcome reporting in submitted manuscripts and compare primary outcomes across published trial protocols, if available, trial registration entries, and submitted manuscripts. To facilitate this process, journals could require authors to submit protocols or trial registrations in addition to trial reports at the time of submission. Peer reviewers could be asked to include in their reviews responses to a series of questions pertaining to the adequacy of outcome reporting. For example, was a single primary outcome specified? If not, were statistical adjustments made for multiple outcomes? In addition, questions could also be asked regarding the consistency of reporting. Was the same primary outcome defined consistently across the various documents? If the primary outcome was inconsistently reported, did the authors provide justification for changing the

primary outcome, and if so, was this justification acceptable [46]? Including these questions in peer review protocols could improve evaluations of the adequacy of outcome reporting and reduce the likelihood that primary outcomes are chosen based on the results of the trial, rather than in advance.

Rules and requirements implemented as policy by journals, however, have generally had only limited effects on the quality of published research and its reporting. Consistent with the findings of the present study, a 2012 Cochrane review reported that journal endorsement of CONSORT was associated with somewhat greater completeness of reporting of published RCT results, but that this association was small [12]. The authors of the review concluded that completeness of reporting continues to be sub-optimal and that fidelity to CONSORT endorsement as an intervention by journals is generally weak. That is, journal editors and reviewers do not consistently enforce policy, even when journals state that they require adherence.

Thus, apart from the possible exception of a handful of highly prestigious scientific and medical journals, it may be unrealistic to think that any single journal will effectively institute meaningful changes without changes in the broader research and publishing cultures, including incentive structures under which research is conducted and disseminated. Clinical trials meaningfully contribute to the advancement of knowledge and improvement of health are those that address important problems with a high level of rigor, regardless of whether outcomes provide evidence for or against intervention effectiveness [47]. Nonetheless, across the spectrum of research production and dissemination, statistically significant or positive findings are often perceived to be more interesting, more important, and more marketable than negative findings and are privileged in publication [47].

It has long been known that when the popular press reports on medical trials to the general public, positive trials are prioritized through more and longer reports and more positive coverage compared to negative trials [48]. Similarly, editors of academic journals have noted that their readership expects exciting trials that show an effect – that is, positive trials. [49]. Additionally, positive studies are cited more than twice as often as negative studies [50], and journal editors are under pressure to maintain or increase their journal's impact factor, which is based on citations [51-52]. Some editors believe that they would put their journals at a competitive disadvantage by rigorously applying trial registration requirements [49]. Even peer reviewers favor trials with statistically significant results in favor of a treatment. In one randomized trial, peer reviewers who thought they were reviewing actual manuscript submissions were substantially more likely to recommend publication of a positive trial over a negative trial and to criticize the methods of the negative trial, even though the positive and negative trial reports had exactly the same methods sections [53]. Among reasons for not publishing negative trial results, researchers have reported that it is not worth the time, that negative results are not important, and that reports of negative results will be rejected by journals [54]. Researchers may also be biased by allegiances to a particular type of intervention, such as in psychotherapy research [55], or by other pressures, such as academic promotion that is dependent on publication in higher-rated journals, which, in turn, may be related to the ability to generate positive results [49].

The present article can contribute to educating the research community about the gravity of the problem of poor trial methodology and reporting. It will not, however, in itself lead to changes that are desperately needed. This will likely require major changes in the culture in which we do research, including re-thinking the very structures by which the research

community interacts with the public and research funders. It may be that funders of research need to take a much more assertive role in determining the criteria by which research funds are allocated, expectations for their use, and the transparency and honesty of their reporting. Currently, researchers themselves are tasked to represent the interests of society through their roles in funding review panels, peer review, and journal editing, but this model has clearly not resulted in the kind of transparent and honest conduct of research that is needed.

There are limitations that should be taken into account when interpreting the results of our study. First, we used a relatively short time span, which resulted in a small number of published articles to review. Second, our study assessed articles published in *Annals of Behavioral Medicine, Health Psychology, Journal of Psychosomatic Research,* and *Psychosomatic Medicine*. The degree to which our results generalize to articles published in other psychosomatic and behavioral health journals is unknown.

In summary, the proportion of registered behavioral health RCTs has markedly increased in the past 5 years. However, the quality of the registry entries and reported outcomes remains largely inadequate. Greater efforts by researchers, peer reviewers, and journal editors are needed to ensure that the identification of the most effective behavioral health interventions is not hindered by reporting biases. More adequate reporting of outcomes and improved quality of trial registrations will ultimately result in less waste in research and better patient care.

References

[1] Devereaux PJ, Yusuf S. The evolution of the randomized controlled trial and its role in evidence-based decision making. J Intern Med. 2003;254:105-13.

[2] Rennie D. CONSORT revised—improving the reporting of randomized trials. JAMA. 2001;285:2006-7.

[3] Schulz KF, Grimes DA. Multiplicity in randomised trials I: endpoints and treatments. Lancet. 2005;365:1591-5.

[4] Moher D, Hopewell S, Schulz KF, Montori V, Gøtzsche PC, Devereaux PJ, et al. CONSORT 2010 explanation and elaboration: updated guidelines for reporting parallel group randomised trials. J Clin Epidemiol. 2010;63:e1-37.

[5] Friedman LM, Furberg C, DeMets DL. Fundamentals of clinical trials. 4th ed. New York: Springer; 2010.

[6] Schulz KF, Grimes DA. Sample size calculations in randomised trials: mandatory and mystical. Lancet. 2005;365:1348-53.

[7] Sterne JAC, Egger M, Moher D. Addressing reporting biases. Cochrane handbook for systematic reviews of interventions: John Wiley & Sons; 2008. p. 297-333.

[8] Dwan K, Altman DG, Cresswell L, Blundell M, Gamble CL, Williamson PR. Comparison of protocols and registry entries to published reports for randomised controlled trials. Cochrane Database Syst Rev. 2011; MR000031.

[9] Page MJ, McKenzie JE, Forbes A. Many scenarios exist for selective inclusion and reporting of results in randomized trials and systematic reviews. J Clin Epidemiol. 2013;66:524-37.

[10] Saquib N, Saquib J, Ioannidis JPA. Practices and impact of primary outcome adjustment in randomized controlled trials: meta-epidemiologic study. BMJ. 2013;347:f4313.

[11] Altman DG, Schulz KF, Moher D, Egger M, Davidoff F, Elbourne D, et al. The revised CONSORT statement for reporting randomized trials: explanation and elaboration. Ann Intern Med. 2001;134:663-94.

[12] Turner L, Shamseer L, Altman D, Schulz K, Moher D. Does use of the CONSORT Statement impact the completeness of reporting of randomised controlled trials published in medical journals? A Cochrane review. Syst Rev. 2012;1:60.

[13] International Committee of Medical Journal Editors. Clinical Trials Registration. www.icmje.org/about-icmje/faqs/clinical-trials-registration/. Accessed February 18, 2015.

[14] Thombs BD, Kwakkenbos L, Coronado-Montoya S. Trial registration in rheumatology: the next step. Arthrit Care Res. 2014;66:1435-7.

[15] De Angelis C, Drazen JM, Frizelle FA, Haug C, Hoey J, Horton R, et al. Is this clinical trial fully registered? A statement from the International Committee of Medical Journal Editors. Can Med Assoc J. 2005;172:1700-2.

[16] Milette K, Roseman M, Thombs BD. Transparency of outcome reporting and trial registration of randomized controlled trials in top psychosomatic and behavioral health journals: A systematic review. J Psychosom Res. 2011;70:205-17.

[17] Kaplan R. Changes in Annals' editorial policies. Ann Behav Med. 2002;24:167-8.

[18] Stone AA. Editorial: Modification to "Instructions to authors". Health Psychol. 2003;22:331.

[19] Sheps DS. Changes in the air at Psychosomatic Medicine. Psychosom Med. 2003;65:499-500.

[20] Creed F, Levenson JL. Raising standards of research reporting in the Journal. J Psychosom Res. 2011;70:203-4.

[21] France CR. Recent changes and future firections for Annals of Behavioral Medicine. Ann Behav Med. 2010;40:1-2.

[22] Victoria White, Managing Editor, Psychosomatic Medicine, Email Communication, July 19th 2012.

[23] Chan AW, Tetzlaff JM, Gøtzsche PC, Altman DG, Mann H, Berlin JA, et al. SPIRIT 2013 explanation and elaboration: guidance for protocols of clinical trials. BMJ. 2013;346:e7586.

[24] Freedland KE, Reese RL, Steinmeyer BC. Multivariable models in biobehavioral research. Psychosom Med. 2009;71:205-16.

[25] Mathieu S, Boutron I, Moher D, Altman DG, Ravaud P. Comparison of registered and published primary outcomes in randomized controlled trials. JAMA. 2009;302:977-84.

[26] Shi Y, Ehlers S, Hinds R, Baumgartner A, Warner DO. Monitoring of exhaled carbon monoxide to promote preoperative smoking abstinence. Health Psychol. 2013;32:714-7.

[27] Sinclair KiA, Makahi EK, Shea-Solatorio C, Yoshimura SR, Townsend CKM, Kaholokula JKa. Outcomes from a diabetes self-management Intervention for Native Hawaiians and Pacific People: Partners in Care. Ann Behav Med. 2013;45:24-32.

[28] John LK, Norton MI. Converging to the lowest common denominator in physical health. Health Psychol. 2013;32:1023-8.

[29] Altman DG, Moher D. Importance of transparent reporting of health research. Guidelines for reporting health research: A user's manual: John Wiley & Sons; 2014. p. 1-13.

[30] Bacon S, Lavoie K, Ninot G, Czajkowski S, Freedland K, Michie S, et al. An international perspective on improving the quality and potential of behavioral clinical trials. Curr Cardiovasc Risk Rep. 2014;9:1-6.

[31] Ioannidis JPA, Greenland S, Hlatky MA, Khoury MJ, Macleod MR, Moher D, et al. Increasing value and reducing waste in research design, conduct, and analysis. Lancet. 2014;383:166-75.

[32] Chan AW, Song F, Vickers A, Jefferson T, Dickersin K, Gøtzsche PC, et al. Increasing value and reducing waste: Addressing inaccessible research. Lancet. 2014;383:257-66.

[33] Begley CG, Ellis LM. Drug development: Raise standards for preclinical cancer research. Nature. 2012;483:531-3.

[34] Perel P, Roberts I, Sena E, Wheble P, Briscoe C, Sandercock P, et al. Comparison of treatment effects between animal experiments and clinical trials: Systematic review. BMJ. 2007;334:197-200.

[35] Prinz F, Schlange T, Asadullah K. Believe it or not: How much can we rely on published data on potential drug targets? Nat Rev Drug Discov. 2011;10:712.

[36] Bracken MB. Why are so many epidemiology associations inflated or wrong? Does poorly conducted animal research suggest implausible hypotheses? Ann Epidemiol. 2009;19:220-4.

[37] Kimmelman J, Anderson JA. Should preclinical studies be registered? Nat Biotech. 2012;30:488-9.

[38] Kimmelman J, Mogil JS, Dirnagl U. Distinguishing between exploratory and confirmatory preclinical research will improve translation. PLoS Biol. 2014;12:e1001863.

[39] Pavlica S. The need for registration of preclinical studies. Medical Writing. 2013;22:131-3.

[40] Smyth RMD, Kirkham JJ, Jacoby A, Altman DG, Gamble C, Williamson PR. Frequency and reasons for outcome reporting bias in clinical trials: Interviews with trialists. BMJ. 2011;342:c7153.

[41] Moher D. Reporting research results: A moral obligation for all researchers. Can J Anesth. 2007;54:331-5.

[42] Hart B, Lundh A, Bero L. Effect of reporting bias on meta-analyses of drug trials: reanalysis of meta-analyses. BMJ. 2012;344:d7202.

[43] Kirkham JJ, Dwan KM, Altman DG, Gamble C, Dodd S, Smyth R, et al. The impact of outcome reporting bias in randomised controlled trials on a cohort of systematic reviews. BMJ. 2010;340:c365.

[44] Hopewell S, McDonald S, Clarke M, Egger M. Grey literature in meta-analyses of randomized trials of health care interventions. Cochrane Database Syst Rev. 2007;MR000010

[45] Higgins JPT, Altman DG, Gøtzsche PC, Jüni P, Moher D, Oxman AD, et al. The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. BMJ. 2011;343:d5928.

[46] Evans S. When and how can endpoints be changed after initiation of a randomized clinical trial. PLOS Clin Trials. 2007;2:e18.

[47] Brice A, Chalmers I. Medical journal editors and publication bias. BMJ. 2013;347:f6170.

[48] Koren G, Klein N. Bias against negative studies in newspaper reports of medical research. JAMA. 1991;266:1824-1826.

[49] Wagner E, Williams P, Overcome Failure to Publish Negative Findings Consortium. "Hardly worth the effort"? Medical journals' policies and their editors' and publishers' views on trial registration and publication bias: quantitative and qualitative study. BMJ. 2013;347:f5248.

[50] Jannot A-S, Agoritsas T, Gayet-Ageron A, Perneger TV. Citation bias favoring statistically significant studies was present in medical research. J Clin Epidemiol. 2013;66:296-301.

[51] Wilhite AW, Fong EA. Coercive citation in academic publishing. Science. 2012;335:542-543.

590 [52] Falagas ME, Alexiou VG. The top-ten in journal impact factor manipulation. Arch Immunol
591 Ther Exp. 2008;56:223-226.
592 [53] Emerson GB, Warme WJ, Wolf FM, Heckman JD, Brand RA, Leopold SS. Testing for the
593 presence of positive-outcome bias in peer review: A randomized controlled trial. Arch Intern
594 Med. 2010;170:1934-9.
595 [54] Song F, Loke Y, Hooper L. Why are medical and health-related studies not being
596 published? A systematic review of reasons given by investigators. PLOS ONE. 2014;9: e110418.
597 [55] Luborsky L, Diguer L, Seligman DA, Rosenthal R, Krause ED, Johnson S, et al. The
598 researcher's own therapy allegiances: A "wild card" in comparisons of treatment efficacy. Clin
599 Psychol Sci Pract. 1999;6:95-106.

Figure 1: Article Selection

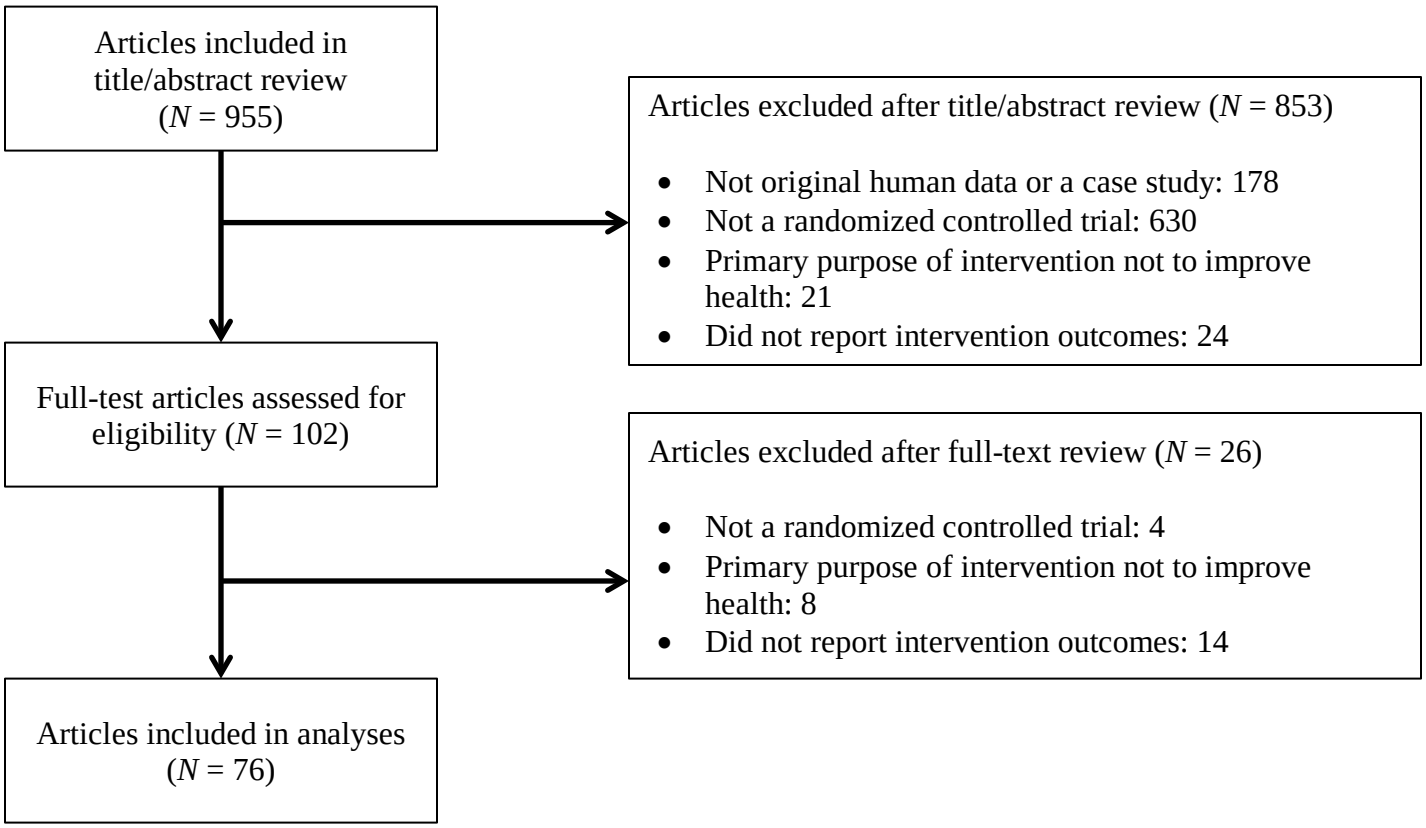

Table I: Outcome Declaration in Published Randomized Controlled Trials

	Annals of Behavioral Medicine	*Health Psychology*	*Journal of Psychosomatic Research*	*Psychosomatic Medicine*	**All Journals (n, % of total)**
Adequately declared outcomes:	**5 (27.8%)**	**13 (39.4%)**	**3 (30.0%)**	**4 (26.7%)**	**25 (32.9%)**
Primary	3	10	2	2	17 (22.4%)
Multiple primary (with statistical adjustment)	1	1	0	0	2 (2.6%)
Secondary	1	2	1	2	6 (7.9%)
Inadequately declared outcomes:	**13 (72.2%)**	**20 (60.6%)**	**7 (70.0%)**	**11 (73.3%)**	**51 (67.1%)**
Multiple primary (same report)	11	10	2	6	29 (38.2%)
Multiple primary (different report)	0	1	1	1	3 (3.9%)
Undefined	2	9	4	4	19 (25.0%)
Total	**18 (100%)**	**33 (100%)**	**10 (100%)**	**15 (100%)**	**76 (100%)**

Table II: Registration of Published RCTs and Pre-Enrolment Registration Requirement per ICMJE Requirements

	Annals of Behavioural Medicine	*Health Psychology*	*Journal of Psychosomatic Research*	*Psychosomatic Medicine*	**All Journals (n, % of total)**
Unregistered RCT Publications	**4 (22.2%)**	**22 (66.7%)**	**4 (40.0%)**	**6 (40.0%)**	**36 (47.4%)**
Registered RCT Publications	**14 (77.8%)**	**11 (33.3%)**	**6 (60.0%)**	**9 (60.0%)**	**40 (52.6%)**
Pre-enrolment registration required	14	11	6	8	39 (51.3%)
Pre-enrolment registration not required	0	0	0	1	1 (1.3%)
Could not assess registration requirement	0	0	0	0	0 (0%)
Total	**18 (100%)**	**33 (100%)**	**10 (100%)**	**15 (100%)**	**76 (100%)**

ICMJE = International Committee of Medical Journal Editors

Appendix

First Author Year	Purpose	Published Outcomes Per Milette et al. [16] Methods: Reported Outcomes	Published Outcomes Per Milette et al. [16] Methods: Adequate / Inadequate Outcome Reporting	Trial Registration: Registration Status	Trial Registration: Registry Name and Number	Trial Registration: Adequate / Inadequate Outcome Registration
Annals of Behavioral Medicine						
Allan 2013	To investigate whether the effectiveness of action plans depends upon the skill of the planner.	Single primary	Adequate	Not registered	Not applicable	Not applicable
Davis 2013	To compare effects of a 12-module online intervention using mindful awareness/acceptance with an attention-control treatment.	Multiple primary (without correction)	Inadequate	Registered*	NCT01748786	Inadequate
Eakin 2013	To test a telephone-delivered weight loss intervention for real-world delivery.	Multiple primary (without correction)	Inadequate	Registered*	ACTRN12608000203358	Inadequate
Jones 2013	To examine the impact of substance use, history of sexual trauma, and intimate partner violence on sexual risk associated with participation in a risk reduction intervention.	Undefined	Inadequate	Not registered	Not applicable	Not applicable
King	To test if sequential versus simultaneous diet plus physical activity	Multiple primary	Inadequate	Registered*	NCT00131105	Inadequate

2013	interventions affects behavior changes.	(without correction)				
Lewis 2013	To examine physical activity behaviors and mediators among adults randomly assigned to either a 6-month print-based theory tailored physical activity intervention or a health/wellness contact control arm.	Multiple primary (without correction)	Inadequate	Registered*	NCT00142688	Inadequate
Loft 2013	To test the efficacy of mental imagery techniques promoting arousal reduction and implementation intentions to improve sleep behavior.	Undefined	Inadequate	Registered*	NCT01648062	Inadequate
McGowan 2013	To determine the effects of an implementation intention intervention on physical activity and quality of life in prostrate cancer survivors.	Multiple primary (without correction)	Inadequate	Registered*	NCT01410656	Inadequate
Morgan 2013	To evaluate the efficacy of two gender-tailored weight-loss interventions for men, which required no face-to-face contact.	Multiple primary (with correction)	Adequate	Registered*	ACTRN12610000699066	Inadequate
Morledge 2013	To determine feasibility of an internet-based mindfulness stress management program.	Multiple primary (without correction)	Inadequate	Registered*	NCT01595555	Inadequate
O'Carroll 2013	To pilot an RCT to improve adherence to prevention medication in stroke survivors using a brief, personalized intervention.	Multiple primary (without correction)	Inadequate	Registered*	ISRCTN38274953	Inadequate
Plotnikoff 2013	To explore the effectiveness of two innovative/theoretically based behavioral change strategies to increase physical activity and reduce hemoglobin A1c in type 2 diabetes adults.	Multiple primary (without correction)	Inadequate	Registered*	NCT00221234	Inadequate
Salmoirago-	To determine the feasibility of a phone-delivered mindfulness intervention in	Secondary	Adequate	Registered*	NCT01035294	Excluded

Blotcher 2013	patients with defibrillators and to obtain preliminary indications of efficacy on mindfulness and anxiety.					(secondary analysis)
Sinclair 2013	To pilot test the effectiveness of a culturally adapted diabetes self-management intervention.	Single primary	Adequate	Registered*	NCT01235429‡	Adequate
Slavin-Spenny 2013	To test an anger awareness and expression intervention and a relaxation intervention compared to each other and waiting list to reduce headache frequency.	Multiple primary (without correction)	Inadequate	Registered*	NCT00956969	Inadequate
Hall 2014	To determine whether implementation intentions are effective for enhancing physical activity in older adult women.	Multiple primary (without correction)	Inadequate	Not registered	Not applicable	Not applicable
Jessop 2014	To explore whether the efficacy of a self-affirmation manipulation at promoting exercise could be enhanced by an implementation intention intervention.	Single primary	Adequate	Not registered	Not applicable	Not applicable
Pakpour 2014	To examine the effects of two message framing interventions on oral self-care behaviors and health among Iranian adolescents.	Multiple primary (without correction)	Inadequate	Registered*	NCT01421108	Inadequate
Health Psychology						
Baldwin 2013	To elucidate causal components related to self-persuasion relevant to making changes in two health behavior domains (physical activity and smoking).	Undefined	Inadequate	Not registered	Not applicable	Not applicable
Bassett-Gunter 2013	To examine the relative effectiveness of chronic disease and psychological health risk information combined with gain- versus loss-framed leisure time physical activity (LTPA) messages for changing perceived personal risk,	Undefined	Inadequate	Not registered	Not applicable	Not applicable

	LTPA response efficacy, and LTPA intentions.					
Carr 2013	To test an internet intervention designed to increase physical activity among adults.	Multiple primary (without correction)	Inadequate	Not registered	Not applicable	Not applicable
Gilliam 2013	To test whether self-presentation as a "good coper" in a public context would improve pain and pain coping.	Undefined	Inadequate	Not registered	Not applicable	Not applicable
Gorin 2013	To evaluate a comprehensive weight-loss program that targeted both an individual's behavior and his or her physical and social home environment.	Multiple primary (without correction)	Inadequate	Registered*	NCT00200330	Inadequate
Hatchell 2013	To test the efficacy of messages based on the Extended Parallel Process Model to increase men's physical activity intentions and behaviors.	Undefined	Inadequate	Not registered	Not applicable	Not applicable
Jensen 2013	To examine the effects of a cognitive-behavioral stress management intervention on indicators of positive psychological well-being and negative psychological well-being in HIV-positive minority women with HPV infection or cervical intraepithelial lesions.	Secondary	Adequate	Not registered	Not applicable	Not applicable
John 2013	To examine how access to information on peer health behaviors affects one's own health behavior.	Multiple primary (without correction)	Inadequate	Registered†	NCT01139541	Adequate
Kerns 2013	To evaluate whether tailored cognitive-behavioral therapy (CBT) that incorporates preferences for learning specific cognitive and/or behavioral skills and uses motivational enhancement strategies would improve treatment engagement and participation compared with standard CBT.	Single primary	Adequate	Registered*	NCT00108381	Inadequate

Lo 2013	To evaluate an intervention based on implementation intention principles designed to increase uptake of colorectal cancer screening in English adults aged 60-69 invited for biennial fecal occult blood testing.	Single primary	Adequate	Not registered	Not applicable	Not applicable
Nyklicek 2013	To examine the effects of a Mindfulness-Based Stress Reduction intervention on cardiovascular and cortisol activity during acute stress.	Undefined	Inadequate	Not registered	Not applicable	Not applicable
Pinto 2013	To evaluate the effects of augmenting oncology health care provider advice with physical activity on health among breast cancer patients.	Single primary	Adequate	Registered*	NCT00230711	Inadequate
Reid 2013	To determine whether personalized normative feedback improves attitudes and practices about sun protection.	Undefined	Inadequate	Not registered	Not applicable	Not applicable
Rivas 2013	To evaluate whether a self-regulation intervention can reduce binge-drinking behavior.	Single primary	Adequate	Not registered	Not applicable	Not applicable
Robinson 2013	To compare the effect on food selection of a message containing health related information about fruit and vegetable consumption with a message containing social normative information about consumption of fruit and vegetables.	Multiple primary (without correction)	Inadequate	Not registered	Not applicable	Not applicable
Safren 2013	To test a prevention intervention for HIV-infected men who have sex with men to reduce HIV sexual transmission risk behavior.	Multiple primary (without correction)	Inadequate	Registered†	NCT00231972	Inadequate
Schuz 2013	To determine whether a self-affirmation based intervention improves risk behavior for skin cancer.	Multiple primary (without correction)	Inadequate	Not registered	Not applicable	Not applicable
Sheeran 2013	To test whether mental contrasting promotes rates of physical activity among overweight, middle-aged, and low-SES men.	Multiple primary (without	Inadequate	Not registered	Not applicable	Not applicable

		correction)				
Shi 2013	To test whether telling patients that their CO levels will be monitored will increase smoking abstinence the morning of surgery.	Single primary	Adequate	Registered*	NCT01014455	Adequate
Watts 2013	To examine the efficacy of an online screening decision aid for men with a family history of prostate cancer.	Single primary	Adequate	Registered*	ACTRN12611000850976	Inadequate
Armistead 2014	To test an HIV preventive intervention among South African youth.	Undefined	Inadequate	Not registered	Not applicable	Not applicable
Breitkopf 2014	To evaluate the effect of a theory-based, culturally targeted intervention on adherence to follow-up among low-income and minority women who experience an abnormal Pap test.	Single primary	Adequate	Registered*	NCT00575510	Inadequate
Cox 2014	To determine the effectiveness of asking Anticipated Regret Questions and risk presentation format on enhancing attitudes and behavioral intentions of mothers to have their daughters receive the HPV vaccine.	Single primary	Adequate	Not registered	Not applicable	Not applicable
Dermen 2014	To provide evidence that, compared with a didactic control intervention, a brief motivational interviewing-based intervention delivered by dental practitioners can yield greater improvements in oral hygiene, health-care utilization, and health outcomes in a population at heightened risk for oral disease.	Single primary	Adequate	Not registered	Not applicable	Not applicable
Dorough 2014	To assess the feasibility and initial efficacy of a primarily electronically delivered intervention for prehypertension.	Undefined	Inadequate	Not registered	Not applicable	Not applicable
Halliwell 2014	To test whether a dissonance intervention is effective when delivered in a school setting to 12- and 13-year-old girls in the United Kingdom.	Undefined	Inadequate	Not registered	Not applicable	Not applicable
Harris	To test whether self-affirmation in the context of a threatening health message	Multiple primary	Inadequate	Not registered	Not applicable	Not applicable

2014	helps promote a health behavior (fruit and vegetable consumption) over a 3-month period, and whether adding an implementation intention enhances the impact of self-affirmation.	(without correction)				
King 2014	To evaluate the sustained 18-month effectiveness on physical activity among inactive midlife to older adults of an automated advisor compared to human counselors.	Multiple primary (different report)	Inadequate	Not registered	Not applicable	Not applicable
Oh 2014	To examine if different exercise intensities acutely reduce snack and cigarette cravings and attentional bias to video clips of snacks and cigarettes among abstinent smokers.	Multiple primary (without correction)	Inadequate	Not registered	Not applicable	Not applicable
Ostroff 2014	To evaluate the efficacy of a novel, pre-surgical cessation intervention in newly diagnosed cancer patients scheduled for surgical hospitalization.	Single primary	Adequate	Registered[†]	NCT00575718	Inadequate
Schwebel 2014	To compare the efficacy of individualized streetside training, training in a virtual pedestrian environment, training using videos and Web sites, plus no-training control, to improve children's street-crossing ability.	Multiple primary (without correction)	Inadequate	Registered[†]	NCT00850759	Inadequate
Sorkin 2014	To evaluate the feasibility of a pilot, dyad-based lifestyle intervention, the Unidas por la Vida program, for improving weight loss and dietary intake among high-risk Mexican American mothers who have Type 2 diabetes and their overweight/obese adult daughters.	Secondary	Adequate	Not registered	Not applicable	Not applicable
Whitehead 2014	To test the efficacy of an appearance-based dietary intervention.	Multiple primary (with correction)	Adequate	Registered[*]	NCT01511484	Inadequate
Journal of Psychosomatic Research						
Björneklett	To perform a long-term follow-up of a support group intervention for women	Undefined	Inadequate	Not registered	Not applicable	Not applicable

2013	after primary breast cancer treatment.					
Brownley 2013	To evaluate whether chromium picolinate is useful in the treatment of binge eating disorder.	Undefined	Inadequate	Registered*	NCT00904306	Inadequate
Fjorback 2013	To conduct a feasibility and efficacy trial of mindfulness therapy in somatization disorder and functional somatic syndromes.	Single primary	Adequate	Registered†	NCT00497185	Inadequate
Hood 2013	To determine the feasibility and initial effects on weight and continuous positive airway pressure of a brief minimal contact self-monitoring-based weight loss intervention.	Single primary	Adequate	Not registered	Not applicable	Not applicable
Probst 2013	To investigate the effects of feedback on psychosomatically treated in-patients at risk of treatment failure.	Undefined	Inadequate	Not registered	Not applicable	Not applicable
Zuidersma 2013	To evaluate the effect of an antidepressant treatment strategy on long-term cardiovascular outcomes and all-cause mortality.	Multiple primary (different report)	Inadequate	Registered*	ISRCTN57865866	Inadequate
Dougall 2014	To compare three therapies (cognitive-behavioral therapy, graded exercise therapy, and adaptive pacing therapy) each added to specialist medical care against specialist medical care alone.	Secondary	Adequate	Registered†	ISRCTN54285094	Excluded (secondary analysis)
Gili 2014	To assess changes in health related quality of life after a cognitive behavioral program for patients diagnosed with abridged somatization disorder in primary care.	Multiple primary (without correction)	Inadequate	Registered†	ISRCTN69944771	Inadequate
Probst 2014	To assess the effect of clinical support tools on patient therapy outcomes.	Undefined	Inadequate	Not registered	Not applicable	Not applicable
van Son 2014	To examine if the effects of a mindfulness-based CBT therapy on emotional distress would be sustained after six month follow-up.	Multiple primary (without	Inadequate	Registered*	Dutch Trial Register NTR2145	Inadequate

		correction)				
Psychosomatic Medicine						
Christian 2013	To test the hypothesis that in contrast to strength training, aerobic training would reduce resting cardiovascular sympathetic indices.	Undefined	Inadequate	Registered*	NCT00358137	Excluded (not required to be registered)
Friedberg 2013	To assess the efficacy of brief fatigue self-management for medically unexplained chronic fatigue and chronic fatigue syndrome in primary care.	Multiple primary (without correction)	Inadequate	Registered*	NCT00997451	Inadequate
Hughes 2013	To determine whether mindfulness-based stress reduction reduces blood pressure compared to progressive muscle relaxation.	Multiple primary (without correction)	Inadequate	Registered*	NCT00440596	Inadequate
Koschwanez 2013	To investigate whether expressive writing could speed wound reepithelialization in healthy, older adults.	Multiple primary (without correction)	Inadequate	Registered*	ACTRN12611000687998	Inadequate
Kotlyar 2013	To examine paroxetine's effect on the physiological response to combining stress and smoking.	Undefined	Inadequate	Registered*	NCT00218439	Inadequate
Long 2013	To assess whether a life-style physical activity intervention improved antibody response to a pneumococcal vaccination in sedentary middle-aged women.	Single primary	Adequate	Not registered	Not applicable	Not applicable
Moreno 2013	To evaluate the effectiveness and feasibility of a cognitive-behavioral program for patients in primary care units who were diagnosed as having abridged somatization disorder.	Multiple primary (without correction)	Inadequate	Registered*	ISRCTN69944771	Inadequate
Pagoto	To demonstrate a scenario in which use of an attention control in a behavioral	Secondary	Adequate	Registered*	NCT00217919	Excluded

2013	trial was unnecessary and possibly detrimental.					(secondary analysis)
Petrowski 2013	To test whether anesthesia versus no anesthesia with pre-hospital discharge testing of an implantable cardioverter-defibrillator is associated with post-discharge anxiety.	Multiple primary (without correction)	Inadequate	Not registered	Not applicable	Not applicable
Vaccarino 2013	To compare the effect of consciously resting meditation to health education on endothelial function in the setting of metabolic syndrome.	Single primary	Adequate	Not registered	Not applicable	Not applicable
Radstaak 2014	To examine whether listening to self-chosen music after stress exposure improves mood, decreases subjective arousal and rumination, and facilitates cardiovascular recovery.	Undefined	Inadequate	Not registered	Not applicable	Not applicable
Rash 2014	To examine the effect of synthetic oxytocin delivered intra-nasally on acute pain sensitivity using a placebo-controlled, double-blind, within-participant crossover design.	Multiple primary (without correction)	Inadequate	Not registered	Not applicable	Not applicable
Stewart 2014	To test the hypothesis that depression treatment delivered before clinical cardiovascular disease onset reduces the risk of cardiovascular disease events.	Multiple primary (different report)	Inadequate	Registered*	NCT01561105	Inadequate
You 2014	To investigate the effects of written emotional disclosure on a model of chronic pain in healthy women with and without trauma history.	Undefined	Inadequate	Not registered	Not applicable	Not applicable
Zernicke 2014	To investigate the feasibility and impact of an online synchronous Mindfulness-Based Cancer Recovery group program for underserved distressed cancer survivors.	Secondary	Adequate	Registered*	NCT01476891	Excluded (secondary analysis)

ACTRN = Australian New Zealand Clinical Trials Registry; ISRCTN = International Standard Randomized Controlled Trial Number Register.
* Registry name and number were reported in the article.
† Registry name and number were not reported in the article.
‡ Registry number reported in the article was incorrect.