

Research waste from poor reporting of core methods and results and redundancy in studies of reporting guideline adherence: a meta-research review

Tiffany Dal Santo, BSc, Master's Student^{1,2}; Danielle B. Rice, PhD, Assistant Professor^{3,4}; Lara S.N. Amiri, Undergraduate Student Intern¹; Amina Tasleem, BA, Research Assistant¹; Kexin Li, BSc, Research Assistant¹; Jill T. Boruff, MLIS, Associate Librarian⁵; Marie-Claude Geoffroy, PhD, Assistant Professor^{2,6,7}; Andrea Benedetti, PhD, Associate Professor⁸⁻¹¹; Brett D. Thombs, PhD, Professor^{1,2,8,9,11,12}

¹Lady Davis Institute for Medical Research, Jewish General Hospital, Montreal, Quebec, Canada; ²Department of Psychiatry, McGill University, Montreal, Quebec, Canada; ³Department of Psychology, St. Joseph's Healthcare Hamilton, Ontario, Canada; ⁴Department of Psychiatry and Behavioural Neurosciences, McMaster University, Ontario, Canada; ⁵Schulich Library of Physical Sciences, Life Sciences, and Engineering, McGill University, Montreal, Quebec, Canada; ⁶Department of Educational and Counselling Psychology, McGill University, Montreal, Quebec, Canada; ⁷McGill Group for Suicide Studies, Douglas Research Centre, Montreal, Quebec, Canada; ⁸Department of Epidemiology, Biostatistics, and Occupational Health, McGill University, Montreal, Quebec, Canada; ⁹Department of Medicine, McGill University, Montreal, Quebec, Canada; ¹⁰Respiratory Epidemiology and Clinical Research Unit, McGill University Health Centre, Montréal, Québec, Canada; ¹¹Biomedical Ethics Unit, McGill University, Montreal, Quebec, Canada; ¹²Department of Psychology, McGill University, Montreal, Quebec, Canada.

Corresponding Author: Brett D. Thombs, PhD; Jewish General Hospital; 3755 Cote Ste

Catherine Road; Montreal, Quebec H3T 1E2; Tel: (514) 340-8222 ext. 24244; e-mail:

brett.thombs@mcgill.ca; ORCID: 0000-0002-5644-8432

Word count: 3,622

ABSTRACT

Objectives: We investigated meta-research studies that evaluated adherence to prominent reporting guidelines (CONSORT, PRISMA, STARD, STROBE) in health research studies to determine the proportion that (1) provided an explanation for how complex guideline items were rated for adherence and (2) provided results from individual studies reviewed in addition to aggregate results. We also examined the conclusions of each meta-research study to assess redundancy of findings across studies.

Design: Cross-sectional meta-research review.

Data sources: MEDLINE (Ovid) searched on July 5, 2022.

Eligibility criteria for selecting studies: Studies in any language were eligible if they used any version of the CONSORT, PRISMA, STARD, or STROBE reporting guidelines or their extensions to evaluate reporting in at least 10 human health research studies. We excluded studies that modified a reporting guideline or its items or evaluated fewer than half of reporting guideline items. Main outcomes were (1) the proportion of meta-research studies that provided a coding explanation that could be used to replicate the study or verify its results and (2) the proportion that provided individual-level study results in the main text, supplemental materials, or via an internet link.

Results: Of 148 included meta-research studies, 14 (10%, 95% confidence interval [CI] 6% to 15%) provided a fully replicable coding explanation, and 49 (33%, 95% CI 26% to 41%) completely reported individual study results. Of 90 studies that classified reporting as adequate or inadequate in the study abstract, 6 (7%, 95% CI 3% to 14%) concluded that reporting was adequate but none of those 6 studies provided information on how items were coded or provided item-level results for included studies.

Conclusions: Much of published meta-research on reporting in health research is likely wasteful.

Few studies report enough information for verification or replication, and almost all find that reporting in health research studies is suboptimal. These findings highlight the importance of shifting the focus from assessing reporting adequacy to developing, testing, and implementing strategies to improve reporting.

Funding: There was no specific funding for this study.

Protocol: Posted on the Open Science Framework June 29, 2022 (<https://osf.io/gtm4z/>).

INTRODUCTION

Research waste occurs when research is not useful because it addresses questions that are already answered; is too poorly designed to answer important research questions; does not make results easily accessible, or reports results too poorly to evaluate or replicate the research.¹⁻⁶ Meta-research studies are conducted to identify areas where research design, conduct, or reporting could be improved and, thus, reduce research waste.⁷ Just like other health research, however, meta-research can be wasteful if it is too poorly designed or reported to be useful or does not add substantively to knowledge.

Many meta-research studies evaluate whether published studies are adequately reported based on reporting guidelines,^{8,9} such as the Consolidated Standards of Reporting Trials (CONSORT),¹⁰ Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA),¹¹ Standards for the Reporting of Diagnostic Accuracy Studies (STARD),¹² or Strengthening the Reporting of Observational Studies in Epidemiology (STROBE).¹³ Reporting guidelines typically include a structured checklist of a minimum set of items that should be reported in studies that use a specific research design along with an explanation document.¹⁴

How meta-researchers translate reporting guideline items into evaluative ratings of individual studies and the ratings generated for included studies are the core elements of any study on reporting guideline adherence. There are no reporting guidelines for meta-research studies, but many of these studies, including studies on reporting guideline adherence, use methods that closely align with systematic review methods.¹⁵⁻²⁰ The PRISMA statement for reporting of systematic reviews and meta-analyses stipulates that methods for collecting and coding data are explicitly defined and that results of each individual study included in a review must be provided.¹¹ Many reporting guideline items are complex and include multiple

components.²¹ Not defining how these items are translated into ratings in meta-research studies creates risk that coding decisions are not made validly or reliably and poses a barrier to replication. Similarly, the failure to report individual study-level results, at least in supplemental material or via a weblink, poses a barrier to reviewing and verifying aggregate findings and does not allow users to identify studies with specific findings that may be of interest.

The objective of our study was to evaluate meta-research studies that reviewed reporting in health research studies using the CONSORT,¹⁰ PRISMA¹¹, STARD,¹² or STROBE¹³ reporting guidelines or one of their extensions to determine the proportion that (1) provided an explanation for how reporting guideline items were translated into adherence ratings with enough information to be replicable and (2) provided results from each individual study included in their report. Additionally, we evaluated the conclusions from the abstract of each study to assess the degree to which meta-research studies on reporting are likely generating new knowledge versus addressing a question for which the answer may already be known.

METHODS

We conducted a cross-sectional evaluation of recently published meta-research studies that evaluated adequacy of health research study reporting based on any version of the CONSORT,¹⁰ PRISMA,¹¹ STARD,¹² or STROBE¹³ reporting guidelines or their extensions. We posted our study protocol on the Open Science Framework (<https://osf.io/gtm4z/>) prior to initiation. Although there are currently no reporting guidelines for meta-research studies, we have reported the present study consistent with applicable PRISMA items as these most closely align with our study design.

Eligibility

Studies published in any language were eligible if they used any version of the CONSORT,¹⁰ PRISMA,¹¹ STARD,¹² or STROBE¹³ reporting guidelines or their extensions (e.g., CONSORT-ROUTINE,²² PRISMA-DTA,²³ STROBE-MR²⁴) to evaluate reporting in human health research publications. To select reporting guidelines for inclusion in our study, we reviewed citations of all EQUATOR guideline publications²¹ and found that these guidelines were by far the most highly cited. To facilitate searching, included studies must have mentioned the name of an eligible guideline in their abstract. Studies that evaluated reporting using multiple reporting guidelines were eligible if at least one of the guidelines was eligible. Studies that investigated reporting as one of multiple research questions or that assessed reporting as part of another research question were eligible. We excluded studies that modified items in an otherwise eligible reporting guideline checklist, added items to an eligible checklist, or evaluated fewer than half of items in an eligible checklist as this could create small subsets of items that might have a different level of coding complexity compared to the full list of checklist items. We excluded studies that evaluated fewer than 10 publications.

Search and Study Selection Method

We searched MEDLINE (ALL) via Ovid for potentially eligible studies using the search strategy: (((quality or complete* or adequat* or transparen*) adj3 reporting) AND (CONSORT* or PRISMA* or STROBE* or STARD* or "Consolidated Standards of Reporting Trials" or "Preferred Reporting Items for Systematic Reviews" or "Standards for Reporting Diagnostic accuracy studies" or "Strengthening the Reporting of Observational Studies in Epidemiology")).tw,kf. The principal investigator (BDT) worked with an experienced health sciences librarian (JTB) to develop the search. The search was then run by a trained research assistant (KL) on July 5, 2022. To include the most recently published meta-research studies and

reflect current practices as best as possible, we reviewed citations identified in the search in reverse chronological order based on their PubMed Unique Identifier until we obtained our targeted sample size of eligible studies. Citations were uploaded to DistillerSR (Evidence Partners, Ottawa, Canada). Two reviewers (TDS, LSNA) independently assessed study eligibility at the title and abstract level. If either reviewer believed a study to be potentially eligible, it moved on to full-text review, where two reviewers (TDS, LSNA, AT) independently assessed eligibility. Discrepancies at the full-text level were resolved by consensus between reviewers with a third reviewer (BDT) consulted as necessary. Supplementary Material 1 includes coding criteria for determining eligibility at the title and abstract and full-text levels.

Sample Size Calculation

A preliminary pre-study review of likely eligible articles published in the 5 years prior to July 2022 found that few studies provided coding definitions or reported individual study results. We therefore hypothesized that the proportion of included articles that provided either would be small. Thus, we set our sample size to have a 95% confidence interval (CI) width of 15% around a percentage reporting of 33%. Based on CIs calculated using the method of Agresti and Coull,²⁵ we sought to obtain 148 studies.

Data Extraction

For each eligible meta-research study, data were extracted in DistillerSR by a single reviewer (TDS, LSNA) and validated by a second reviewer (TDS, LSNA, AT) using the DistillerSR Quality Control function. Discrepancies were resolved by consensus between reviewers with a third reviewer (BDT) consulted as necessary. The data extraction form can be found in Supplementary Material 2. Reviewers extracted (1) publication characteristics (first author last name; publication year; journal and 2021 impact factor); (2) corresponding author

country; (3) research question (reporting was the only research question; reporting was the main research question and there were other non-reporting questions; there were multiple research questions, including reporting and non-reporting questions, and the main one is unclear; the main research question was not reporting, but an eligible reporting analysis was conducted) (4) reporting guideline(s) evaluated; (5) number of publications included in the study; (6) main eligibility criteria of included publications (by reporting guideline, study design, field of research, patient population, intervention type, journal, other); (7) number and independence of raters and the rating method used (e.g., yes/no, fully/partially/not reported); and (8) conclusion about reporting adequacy. We used publication abstracts to extract conclusions as these are the most read, and in many cases, the only part of an article that is read.^{26,27} If a study's supplementary material was not accessible via the publishing journal's website, we contacted the corresponding author and the journal editorial manager or editor-in-chief to request access. We sent up to 2 follow-up emails to corresponding authors and journal staff until access was obtained; if we did not receive a response, we coded the study based on available information because no study user would have access to the supplementary material.

To answer our main research questions, reviewers extracted (1) whether the authors provided coding explanations for translating items into ratings with enough information to be replicated and (2) if the authors provided results for each individual study included in their report. Authors may have provided coding explanations and individual study results in the main text or table of their study, in supplementary material, or via an internet link. Coding explanations must have specifically reported which parts of each item were required for the item to be used in evaluating reporting adequacy. For individual study results, we coded whether authors reported results for each item for all studies, reported results partially (e.g., an overall

score but not item ratings for each study), or did not report individual study results. Our coding manual for evaluating coding explanations and reporting of individual study results are found in Supplementary Material 3 and 4.

Analysis

We calculated the (1) proportion of meta-research studies that provided a coding guide for translating reporting guideline items into ratings with enough information to be replicated and (2) proportion that provided results for each individual study included in their report. All proportions are presented with 95% CIs using the method of Agresti and Coull.²⁵ We also presented results by subgroups defined by corresponding author country, 2021 journal impact factor, reporting guideline evaluated (CONSORT, PRISMA, STROBE, STARD), and research question (reporting is part of the main research question; reporting is not part of the main research question). When presenting outcomes by subgroups, we included guideline extensions (e.g., CONSORT-ROUTINE) with the main guideline (e.g., CONSORT). Groupings (e.g., impact factor levels) were established based on available frequency data prior to examining reporting results. We did not conduct statistical tests to compare subgroups because our study was not designed or powered for that purpose.

Patient and Public Involvement

Patients and members of the public were not involved in the design, conduct, reporting, or dissemination plans of our research.

RESULTS

Search Results and Included Study Characteristics

Our search yielded 1,698 unique titles and abstracts. We reviewed these in reverse chronological order until we obtained 148 included studies. During this process we reviewed 418

titles and abstracts. We excluded 182 studies at the title and abstract level, and 88 studies at the full-text level (Figure 1).

The 148 included studies were initially listed on PubMed between August 14, 2020 and June 30, 2022. They assessed between 10 and 2,844 studies (median = 52; interquartile range = 24 to 120). As shown in Table 1, included studies were from China (N = 51; 34%), the United States (N = 27; 18%), the United Kingdom (N = 9; 6%), Canada (N = 8; 5%), and 22 other countries that had 1 to 5 included studies (N = 53; 36%). Most studies assessed adherence to the CONSORT (N = 61; 41%) or PRISMA (N = 59; 40%) reporting guidelines or their extensions. Reporting was the only research question in 46 (31%) studies, the main question among multiple questions in 13 (9%) studies, one of multiple questions with no clear primary question in 65 (44%) studies, and not part of the main question in 24 (16%) studies. Most studies 103 (70%) came from journals with impact factor > 2.9. See Supplementary Material 5 for individual study characteristics.

Of the 148 included studies, 3 studies (2%, 95% CI 1% to 6%) used 1 rater, 10 (7%, 95% CI 4% to 12%) used 1 rater with validation from a second rater, 113 (76%, 95% CI 69% to 83%) used 2 or more independent raters, 9 (6%, 95% CI 3% to 11%) used 2 or more raters but did not state whether they were independent, 3 (2%, 95% CI 1% to 6%) used other methods, and 10 (7%, 95% CI 4% to 12%) did not report how many raters were used. For classifying adherence to reporting checklist items, 66 studies (45%, 95% CI 37% to 53%) classified items dichotomously, 61 (41%, 95% CI 34% to 49%) used a multi-level approach (e.g., coded item as “fully reported”, “partially reported”, or “not reported”), 2 (1%, 95% CI 0% to 5%) classified some items dichotomously and others with a multi-level approach, and 19 (13%, 95% CI 8% to 19%) did not report how they classified items. See Supplementary Material 6.

Main Outcomes

Of the 148 studies, 14 (10%, 95% CI 6% to 15%) provided a fully replicable coding explanation, 5 (3%, 95% CI 2% to 8%) provided a partially replicable coding explanation, and 129 (87%, 95% CI 81% to 92%) did not provide enough information to know how coding decisions had been made (see Table 2). Forty-nine studies (33%, 95% CI 26% to 41%) completely reported individual study results, 26 (18%, 95% CI 12% to 25%) reported partial results for all studies, 3 (2%, 95% CI 1% to 6%) reported results for some studies but not others, and 70 (47%, 95% CI 39% to 55%) did not provide any individual study results (see Table 3). Only 4 (3%, 95% CI 1% to 7%) studies provided both fully replicable coding explanations and completely reported individual study results.

One hundred and twenty-two studies mentioned reporting in their abstract conclusion, and 90 of these classified reporting as either adequate or inadequate. Of these 90 studies, 6 (7%, 95% CI 3% to 14%) concluded that reporting was adequate, 29 (32%, 95% CI 24% to 42%) implicitly concluded that reporting was inadequate, and 55 (61%, 95% CI 51% to 71%) explicitly concluded that reporting was inadequate. Of the 6 studies that concluded that reporting was adequate, none provided any explanation of how items were coded, and none provided item-level results for individual studies. The 4 studies with fully replicable coding explanations and complete individual study results all concluded that reporting of individual studies was inadequate. See Table 4. Outcomes for all individual meta-research studies we reviewed are shown in Supplementary Material 7.

As shown in Tables 2, 3 and 4, most subgroup results did not differ substantively from overall conclusions, excluding subgroups with very small numbers of meta-research studies (e.g., < 10 studies). One exception was for subgroups by the research question. Of the 124 studies

where reporting was part of the main research questions, 35 (28%, 95% CI 21% to 37%) completely reported individual study results, whereas 14 of 24 studies (58%, 95% CI 39% to 76%) where reporting was not the main research question reported individual study results. These 24 studies were mostly systematic reviews with clinical questions that assessed reporting as part of their evaluation of included studies.

DISCUSSION

We examined 148 studies that evaluated reporting guideline adherence across areas of health research and were initially listed in PubMed between August 2020 and June 2022. Of the 148 studies, only 10% provided enough information on how individual checklist items were rated to allow for verification of how study methods had been applied or for others to replicate them. Only 33% reported results for all studies they evaluated, at least in supplementary material or via an internet link, so that results could be verified. We did not identify any substantive differences by subgroups such as country or impact factor of the journal where studies were published. Of 90 studies that classified reporting as adequate or inadequate in their abstracts, 7% concluded that reporting was adequate; however, none of the studies that described reporting as adequate provided an explanation of how they coded items or provided item-level results for individual studies. Only 3% of included meta-research studies provided both fully replicable coding explanations and complete individual study results, and all of those studies concluded that reporting was inadequate.

No previous studies have examined the degree to which meta-research studies on reporting guideline adherence adequately report key aspects of their own studies. Given that meta-research is done to scrutinize research methodology and conduct and to promote methods that support the validity and integrity of science,²⁸ an implicit assumption is that meta-research

studies themselves are rigorously conducted and reported. Our study shows that this is often not the case. Dozens of reviews of studies on health research reporting have underlined that reporting is consistently poor.²⁹ Our findings show that many of the meta-research studies that have led to those conclusions are themselves poorly reported.

Good research asks important questions and uses methods that allow us to be confident in its conclusions.³⁰ Large numbers of studies are being done on reporting adherence, virtually all finding the same thing, but few of them are well-enough conducted and reported for more nuanced conclusions to be drawn with confidence. Researchers who are considering initiating a study on adherence to reporting guidelines and editors who must decide whether to publish such studies should be able to clearly articulate the specific problem they are addressing and how they might add to what we already know about the state of research reporting. For example, evaluating reporting to understand the influence of new or modified reporting guidelines or to assess the effects of interventions designed to improve reporting would likely be justified. Simply documenting poor reporting guideline adherence in one more sub-specialty area would likely not be useful. Authors of any studies that do evaluate reporting should clearly describe how reporting was evaluated and should provide study-level information so that others can evaluate and validate their findings. Reporting guidelines for meta-research studies do not yet exist, but a protocol for such guidelines has been published.¹⁶ The authors of these proposed guidelines should ensure that meta-research studies on reporting, in addition to other important items, address the reporting gaps we have identified here.

Rather than more studies on the poor quality of reporting in health care research, interventions that help researchers, peer reviewers, and journal editors improve reporting are needed. A 2019 scoping review identified 31 interventions created to improve reporting

guideline adherence but only 11 that had been evaluated in any way.³¹ Strategies varied on what step of the writing or publishing process they targeted, but most have been aimed at improving adherence at the journal level, such as editorial endorsement of specific reporting guidelines, instructions to authors that recommend or require adherence to reporting guidelines, or requiring authors to submit a completed reporting checklist. The scoping review found only 4 randomised trials of interventions to enhance adherence, but the only one that showed a statistically significant effect of the intervention was the Consort-based WEB (COBWEB) tool, that supports adherence at the writing stage of the manuscript.³² The tool divides CONSORT items into bullet points and emphasizes key reporting elements that need to be reported for the main CONSORT checklist and selected extensions.³² In the trial of the COBWEB tool, which included 41 participants, the global score for completeness of reporting (0-10 scale) was 2.1 points higher (95% CI 1.5 to 2.7) in 123 CONSORT domains drafted with the tool compared to 123 domains drafted without the use of the tool.³² Another intervention, which was published after the search period of the scoping review, in which a journal required authors to incorporate section headings that reflected CONSORT items into their manuscripts, also improved reporting.³³ Overall, however, there are few interventions that have been tested in randomised trials and found to be effective. Resources should be allocated to developing, testing, and disseminating effective interventions that address different aspects of the complex factors that contribute to how well research is reported.³¹

Strengths and Limitations

Strengths of our study include that it is the first study, to the best of our knowledge, to assess reporting in meta-research studies on research reporting and, thus, addresses an important and largely unrecognized problem. We developed and posted a protocol prior to initiating the

study, and we have provided all our coding manuals and individual study results in supplementary material. We included a large sample size of the most recently published studies based on an a priori power analysis.

There are some limitations that also need to be considered. First, we only searched MEDLINE, which could have led us to miss potentially eligible studies, although we found that reporting was poor, and it is unlikely that health research studies in other databases, but not MEDLINE, would have been more completely reported. Second, we included meta-research studies that assessed adherence to 4 EQUATOR reporting guidelines based on how often they have been cited, but we did not assess other reporting guidelines. We do not believe that including other reporting guidelines would have influenced results substantively considering that we assessed reporting in the meta-research studies themselves and not reporting levels of studies that used those reporting guidelines.

CONCLUSIONS

To conclude, we found that out of the 148 studies we assessed, 10% provided a fully replicable coding explanation, 33% completely reported individual study results, and 7% of those that categorized reporting as being adequate or inadequate concluded that adherence to reporting guidelines was adequate, though none of the studies that rated reporting as adequate were themselves well reported. Meta-research is done to reduce research waste by improving how research is performed, communicated, and used,²⁸ but our study shows that meta-research on reporting may be a significant contributor to waste. Most recent studies on reporting guideline adherence do not appear to have added meaningfully to what we know about the problem of research reporting. Poor reporting of key elements in most of these studies does now allow us to draw conclusions beyond that overall reporting continues to be sub-optimal, such as

understanding where reporting gaps are most salient or how to address them. New studies on adherence should only be conducted if there is a specific and justified rationale to address a well-defined, non-redundant research question, such as whether implementation of an intervention changes reporting practices. Rather than more research on poor reporting in yet another sub-specialty area, research is needed that develops effective interventions to improve reporting, tests them in randomised trials, and disseminates them via support and training tools.

FIGURE LEGEND

Figure 1. PRISMA flow diagram.

Table 1: Study characteristics (N = 148)

Study Characteristics	N (%)
Year Published	
2020	21 (14)
2021	60 (41)
2022	56 (38)
Online only	11 (7)
Country	
Canada	8 (5)
China	51 (34)
United Kingdom	9 (6)
United States	27 (18)
Other (all with ≤ 5 studies) ^a	53 (36)
Journal Impact Factor^b	
Impact factor ≤ 2.9	45 (30)
Impact factor > 2.9	103 (70)
Included Study Eligibility Criteria^c	
Study design	137 (93)
Patient population	68 (46)
Intervention type	65 (44)
Journal	17 (11)
Included in specified guidelines	13 (9)
Field of research	13 (9)
Other ^d	6 (4)
Research Question	
Reporting is only research question	46 (31)
Reporting is main of multiple research questions	13 (9)
Multiple research questions with main unclear	65 (44)
Main research question not reporting	24 (16)
Reporting Guideline^e	
CONSORT	61 (41)
PRISMA	59 (40)
STARD	10 (7)
STROBE	18 (12)
Number of Included Publications Reviewed	
≤ 50	72 (49)
>50	76 (51)

It is made available under a [CC-BY-NC 4.0 International license](#) .

^aAustralia (3); Brazil (3); Chile (1); Croatia (1); France (3); Germany (4); Greece (2); India (3); Iran (2); Ireland (2); Italy (3); Korea (4); Macao (2); Mexico (1); Portugal (1); Qatar (2); Saudi Arabia (2); South Africa (1); South Korea (4); Spain (3); Switzerland (1); the Netherlands (5). ^bJournals for which we could not find an impact factor were coded as 0. ^cIncluded reviews could be counted in more than one category. ^dStudies reviewed included a specific questionnaire, were on acceptability of a specific intervention, were abstracts submitted to specific conferences, or were studies that used a specific database . ^eIncluding extensions to specified reporting guidelines.

Table 2: Number and percent of studies that provided fully or partially replicable coding explanations or did not provide coding explanations for overall sample (N = 148) and subgroups

Subgroups	N % (95%CI)		
	Fully Replicable	Partially Replicable	Not Replicable
All	14 10% (6%, 15%)	5 3% (2%, 8%)	129 87% (81%, 92%)
Country			
Canada	3 38% (14%, 69%)	2 25% (7%, 59%)	3 38% (14%, 69%)
China	3 6% (2%, 16%)	0 0% (0%, 7%)	48 94% (84%, 98%)
United Kingdom	1 11% (2%, 44%)	1 11% (2%, 44%)	7 78% (45%, 94%)
United States	2 7% (2%, 23%)	0 0% (0%, 13%)	25 93% (77%, 98%)
Other	5 9% (4%, 20%)	2 4% (1%, 13%)	46 87% (75%, 94%)
Journal Impact Factor			
Impact factor \leq 2.9	1 2% (0%, 12%)	0 0% (0%, 8%)	44 98% (88%, 100%)
Impact factor > 2.9	13 13% (8%, 20%)	5 5% (2%, 11%)	85 83% (74%, 89%)
Reporting Guideline			
CONSORT & extensions	9 15% (8%, 26%)	3 5% (2%, 14%)	49 80% (69%, 88%)
PRISMA & extensions	1 2% (0%, 9%)	1 2% (0%, 9%)	57 97% (89%, 99%)
STARD & extensions	2 20% (6%, 51%)	1 10% (2%, 40%)	7 70% (40%, 89%)

STROBE & extensions	2 11% (3%, 33%)	0 0% (0%, 18%)	16 89% (67%, 97%)
---------------------	--------------------	-------------------	----------------------

Research Question

Reporting is the only research question or there are multiple research questions and the main one is reporting or not defined	14 11% (7%, 18%)	4 3% (1%, 8%)	106 86% (78%, 91%)
Reporting is not the main research question	0 0% (0%, 14%)	1 4% (1%, 20%)	23 96% (80%, 99%)

Table 3: Level of reporting of included study results for overall sample (N = 148) and subgroups

Subgroups	N % (95%CI)			
	Completely Reported	Partially Reported – All Studies	Partially Reported – Some Studies	Not Reported
All	49 33 % (26%, 41%)	26 18% (12%, 25%)	3 2% (1%, 6%)	70 47 % (39%, 55%)
Country				
Canada	2 25% (7%, 59%)	0 0% (0%, 32%)	0 0% (0%, 32%)	6 75% (41%, 93%)
China	26 51% (38%, 64%)	2 4% (1%, 13%)	0 0% (0%, 7%)	23 45% (32%, 59%)
United Kingdom	4 44% (19%, 73%)	2 22% (6%, 55%)	0 0% (0%, 30%)	3 33% (12%, 65%)
United States	3 11% (4%, 28%)	13 48% (31%, 66%)	1 4% (1%, 18%)	10 37% (22%, 56%)
Other	14 26% (16%, 40%)	9 17% (9%, 29%)	2 4% (1% 13%)	28 53% (40%, 66%)
Journal Impact Factor				
Impact factor ≤ 2.9	16 36% (23%, 50%)	9 20% (11%, 34%)	1 2% (0%, 12%)	19 42% (29%, 57%)
Impact factor > 2.9	33 32% (24%, 42%)	17 17% (11%, 25%)	2 2% (1%, 7%)	51 50% (40%, 59%)
Reporting Guideline				
CONSORT & extensions	14 23% (14%, 35%)	6 10% (5%, 20%)	0 0% (0%, 6%)	41 67% (55%, 78%)
PRISMA & extensions	25 42% (31%, 55%)	15 25% (16%, 38%)	1 2% (0%, 9%)	18 31% (20%, 43%)
STARD & extensions	3 30% (11%, 60%)	1 10% (2%, 40%)	1 10% (2%, 40%)	5 50% (24%, 76%)

STROBE & extensions	7 39% (20%, 61%)	4 22% (9%, 45%)	1 6% (1%, 26%)	6 33% (16%, 56%)
---------------------	---------------------	--------------------	-------------------	---------------------

Research Question

Reporting is the only research question or there are multiple research questions and the main one is reporting or not defined	35 28% (21%, 37%)	22 18% (12%, 25%)	2 2% (0%, 6%)	65 52% (44%, 61%)
Reporting is not the main research question	14 58% (39%, 76%)	4 17% (7%, 36%)	1 4% (1%, 20%)	5 21% (9%, 41%)

Table 4: Conclusions in abstracts of included studies on research reporting for overall sample (N = 148) and subgroups

Subgroups	N % (95%CI)					
	Adequate	Inadequate - Implicit	Inadequate - Explicit	Mixed	Vague	No Mention
All	6 4% (2%, 9%)	29 20% (14%, 27%)	55 37% (30%, 45%)	10 7% (4%, 12%)	22 15% (10%, 22%)	26 18% (12%, 25%)
Country						
Canada	0 0% (0%, 32%)	0 0% (0%, 32%)	2 25% (7%, 59%)	1 13% (2%, 47%)	3 38% (14%, 69%)	2 25% (7%, 59%)
China	1 2% (0%, 10%)	15 29% (19%, 43%)	21 41% (29%, 55%)	0 0% (0%, 7%)	6 12% (6%, 23%)	8 16% (8%, 28%)
United Kingdom	0 0% (0%, 30%)	2 22% (6%, 55%)	3 33% (12%, 65%)	1 11% (2%, 44%)	1 11% (2%, 44%)	2 22% (6%, 55%)
United States	1 4% (1%, 18%)	2 7% (2%, 23%)	12 44% (28%, 63%)	4 15% (6%, 33%)	4 15% (6%, 33%)	4 15% (6%, 33%)
Other	4 8% (3%, 18%)	10 19% (11%, 31%)	17 32% (21%, 46%)	4 8% (3%, 18%)	8 15% (8%, 27%)	10 19% (11%, 31%)
Journal Impact Factor						
Impact factor ≤ 2.9	2 4% (1%, 15%)	19 42% (29%, 57%)	9 20% (11%, 34%)	4 9% (4%, 21%)	7 16% (8%, 29%)	4 9% (4%, 21%)
Impact factor > 2.9	4 4% (2%, 10%)	36 35% (26%, 45%)	20 19% (13%, 28%)	6 6% (3%, 12%)	15 15% (9%, 23%)	22 21% (15%, 30%)
Reporting Guidelines						
CONSORT & extensions	3 5% (2%, 14%)	12 20% (12%, 31%)	29 48% (36%, 60%)	3 5% (2%, 14%)	9 15% (8%, 26%)	5 8% (4%, 18%)
PRISMA & extensions	1 2% (0%, 9%)	15 25% (16%, 38%)	16 27% (17%, 40%)	7 12% (6%, 23%)	11 19% (11%, 30%)	9 15% (8%, 27%)
STARD & extensions	0	2	4	0	2	2

STROBE & extensions	0% (0%, 28%) 2 11% (3%, 33%)	20% (6%, 51%) 0 0% (0%, 18%)	40% (17%, 69%) 6 33% (16%, 56%)	0% (0%, 28%) 0 0% (0%, 18%)	20% (6%, 51%) 0 0% (0%, 18%)	20% (6%, 51%) 10 56% (34%, 75%)
---------------------	------------------------------------	------------------------------------	---------------------------------------	-----------------------------------	------------------------------------	---------------------------------------

Research Question

Reporting is the only research question or there are multiple research questions and the main one is reporting or not defined	6 5% (2%, 10%)	54 44% (35%, 52%)	25 20% (14%, 28%)	10 8% (4%, 14%)	20 16% (11%, 24%)	9 7% (4%, 13%)
Reporting is not the main research question	0 0% (0%, 14%)	1 4% (1%, 20%)	4 17% (7%, 36%)	0 0% (0%, 14%)	2 8% (2%, 26%)	17 71% (51%, 85%)

Contributions: TDS, DBR, MCG, AB, and BDT were responsible for the study conception and design. JTB and BDT were responsible for the design of the database search. KL carried out the search. TDS, LSNA, AT, and BDT, contributed to data extraction, coding, and evaluation of included studies. TDS conducted the analyses. TDS drafted the manuscript, and DBR, LSNA, AT, KL, JTB, MCG, AB, and BDT provided critical review and approved the final manuscript.

Copyright for Authors: The Corresponding Author has the right to grant on behalf of all authors and does grant on behalf of all authors, to the Publishers and its licensees in perpetuity, in all forms, formats and media (whether known now or created in the future), to i) publish, reproduce, distribute, display and store the Contribution, ii) translate the Contribution into other languages, create adaptations, reprints, include within collections and create summaries, extracts and/or, abstracts of the Contribution, iii) create any other derivative work(s) based on the Contribution, iv) to exploit all subsidiary rights in the Contribution, v) the inclusion of electronic links from the Contribution to third party material where-ever it may be located; and, vi) licence any third party to do any or all of the above.

The Corresponding Author has the right to grant on behalf of all authors and does grant on behalf of all authors, an exclusive licence (or non-exclusive for government employees) on a worldwide basis to the BMJ Publishing Group Ltd to permit this article (if accepted) to be published in BMJ editions and any other BMJ PGL products and sublicences such use and exploit all subsidiary rights, as set out in our licence.

Funding: There was no specific funding for this study. TDS was supported by a Canadian Institutes of Health Research Masters Award, LSNA by a Mitacs Globalink internship, MCG (Tier 2) and BDT (Tier 1) by Canada Research Chairs, and AB by a Fonds de recherche du Québec – Santé senior researcher salary award, all outside of the present work.

Declaration of Competing Interests: All authors have completed the ICJME uniform disclosure form at www.icmje.org/coi_disclosure.pdf and declare: no support from any organisation for the submitted work; no financial relationships with any organisations that might have an interest in the submitted work in the previous three years. All authors declare no relationships or activities that could appear to have influenced the submitted work. No funder had any role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the manuscript; and decision to submit the manuscript for publication. DBR and BDT declared that they were named or group authors of 3 included studies^{S100,S111,S112} conducted to benchmark reporting prior to publishing a new reporting guideline.

Ethical Approval: As this study was a meta-research study of published results, it did not require ethical approval.

Transparency Declaration: The manuscript's guarantor affirms that the manuscript is an honest, accurate, and transparent account of the study being reported; that no important aspects of the study have been omitted; and that any discrepancies from the study as planned (and, if relevant, registered) have been explained.

Data Sharing: All data used in the study are available in the manuscript and its tables or supplementary files.

Dissemination to study participants or patient communities: There are no plans to disseminate the results of the research directly to a relevant patient community.

Provenance and peer review: Not commissioned; externally peer reviewed.

CC BY licence: The default licence, a CC BY NC licence, is needed.

What is already known on this topic:

- Many meta-research studies have reported that most health research does not adhere to reporting guidelines.
- Meta-research studies are presumably conducted rigorously and fill an important role of reducing waste by identifying areas where we can improve research practice; no studies have evaluated the degree to which meta-research studies on reporting are themselves well-enough reported and add meaningfully to knowledge to fill this role.

What this study adds:

- We reviewed 148 meta-research studies published in 2020 to 2022 that evaluated reporting guideline adherence in human health research.
- We found that only 10% of meta-research studies included enough information on their data coding methods to verify results or replicate the studies, only 33% provided results for their coding of individual included studies, and almost all reached the same conclusion that reporting is not adequate.
- Our findings demonstrate that much meta-research on reporting guideline adherence likely contributes to research waste due to poor reporting and redundancy, and we need to shift our focus to developing, testing, and disseminating effective strategies to improve reporting.

REFERENCES

1. Chalmers I, Bracken MB, Djulbegovic B, et al. How to increase value and reduce waste when research priorities are set. *Lancet* 2014;383:156-165. doi:10.1016/S0140-6736(13)62229-1
2. Chalmers I, Glasziou P. Avoidable waste in the production and reporting of research evidence. *Lancet* 2009;374:86-89. doi:10.1016/S0140-6736(09)60329-9
3. Glasziou P, Chalmers I. Research waste is still a scandal—an essay by Paul Glasziou and Iain Chalmers. *BMJ* 2018;363:k4645. doi:10.1136/bmj.k4645
4. Glasziou P, Altman DG, Bossuyt P, et al. Reducing waste from incomplete or unusable reports of biomedical research. *Lancet* 2014;383:267-276. doi:10.1016/S0140-6736(13)62228-X
5. Grainger MJ, Bolam FC, Stewart GB, et al. Evidence synthesis for tackling research waste. *Nat Ecol Evol* 2020;4:495-497. doi:10.1038/s41559-020-1141-6
6. Macleod MR, Michie S, Roberts I, et al. Biomedical research: increasing value, reducing waste. *Lancet* 2014;383:101-104. doi:10.1016/S0140-6736(13)62329-6
7. Ioannidis JPA, Fanelli D, Dunne DD, et al. Meta-research: evaluation and improvement of research methods and practices. *PLoS Biol* 2015;13:e1002264. doi:10.1371/journal.pbio.1002264

8. Jin Y, Sanger N, Shams I, et al. Does the medical literature remain inadequately described despite having reporting guidelines for 21 years? - A systematic review of reviews: an update. *J Multidiscip Healthc* 2018;11:495-510. doi:10.2147/JMDH.S155103
9. Pussegoda K, Turner L, Garritty C, et al. Systematic review adherence to methodological or reporting quality. *Syst Rev* 2017;6:131. doi:10.1186/s13643-017-0527-2
10. Schulz KF, Altman DG, Moher D, et al. CONSORT 2010 statement: updated guidelines for reporting parallel group randomised trials. *BMC Med* 2010;8:18. doi:10.1186/1741-7015-8-
11. Page MJ, McKenzie JE, Bossuyt PM, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ* 2021;372:n71. doi:10.1136/bmj.n71
12. Bossuyt PM, Reitsma JB, Bruns DE, et al. STARD 2015: an updated list of essential items for reporting diagnostic accuracy studies. *BMJ* 2015;351:h5527. doi:10.1136/bmj.h5527
13. von Elm E, Altman DG, Egger M, et al. Strengthening the reporting of observational studies in epidemiology (STROBE) statement: guidelines for reporting observational studies. *BMJ* 2007;335:806-808. doi:10.1136/bmj.39335.541782.AD
14. The EQUATOR Network. What is a reporting guideline? Accessed December 12, 2022. <https://www.equator-network.org/about-us/what-is-a-reporting-guideline/#:~:text=A%20reporting%20guideline%20provides%20a,Included%20in%20a%20systematic%20review.>
15. Murad MH, Wang Z. Guidelines for reporting meta-epidemiological methodology research. *Evid Based Med* 2017;22:139-142. doi:10.1136/ebmed-2017-110713

16. Lawson DO, Puljak L, Pieper D, et al. Reporting of methodological studies in health research: a protocol for the development of the MethodologIcal STudy reportIng Checklist (MISTIC). *BMJ Open* 2020;10:e040478. doi:10.1136/bmjopen-2020-040478
17. Zhang Z. Meta-epidemiological study: a step-by-step approach by using R. *J Evid Based Med* 2016;9:91-97. doi:10.1111/jebm.12191
18. Gundogan B, Dowlut N, Rajmohan S, et al. Assessing the compliance of systematic review articles published in leading dermatology journals with the PRISMA statement guidelines: a systematic review. *JAAD Int* 2020;1:157-174. doi:10.1016/j.jdin.2020.07.007
19. Ritchie A, Seubert L, Clifford R, et al. Do randomised controlled trials relevant to pharmacy meet best practice standards for quality conduct and reporting? A systematic review. *Int J Pharm Pract* 2020;28:220-232. doi:10.1111/ijpp.12578
20. Canagarajah NA, Porter GJ, Mitra K, et al. Reporting quality of surgical randomised controlled trials in head and neck cancer: a systematic review. *Eur Arch Otorhinolaryngol* 2021;278:4125-4133. doi:10.1007/s00405-021-06694-9
21. The EQUATOR Network. Enhancing the QUALity and Transparency Of health Research. Accessed December 12, 2022. <https://www.equator-network.org/>
22. Kwakkenbos L, Imran M, McCall SJ, et al. CONSORT extension for the reporting of randomised controlled trials conducted using cohorts and routinely collected data (CONSORT-ROUTINE): checklist with explanation and elaboration. *BMJ* 2021;373:n857. doi:10.1136/bmj.n857

23. McInnes MDF, Moher D, Thoms BD, et al. Preferred reporting items for a systematic review and meta-analysis of diagnostic test accuracy studies: the PRISMA-DTA statement. *JAMA*. 2018;319:388-396. doi:10.1001/jama.2017.19163
24. Skrivankova VW, Richmond RC, Woolf BAR, et al. Strengthening the reporting of observational studies in epidemiology using mendelian randomization: the STROBE-MR statement. *JAMA* 2021;326:1614-1621. doi:10.1001/jama.2021.18236
25. Agresti A, Coull BA. Approximate is better than “exact” for interval estimation of binomial proportions. *Am Stat* 1998;52:119-126. doi:10.2307/2685469
26. Beller EM, Glasziou PP, Altman DG, et al. PRISMA for Abstracts: reporting systematic reviews in journal and conference abstracts. *PLoS Med* 2013;10:e1001419. doi:10.1371/journal.pmed.1001419
27. Pitkin RM, Branagan MA. Can the accuracy of abstracts be improved by providing specific instructions? A randomized controlled trial. *JAMA* 1998;280:267-269. doi:10.1001/jama.280.3.267
28. Ioannidis JPA. Meta-research: why research on research matters. *PLoS Biol* 2018;16:e2005468. doi:10.1371/journal.pbio.2005468
29. Samaan Z, Mbuagbaw L, Kosa D, et al. A systematic scoping review of adherence to reporting guidelines in health care literature. *J Multidiscip Healthc* 2013;6:169-188. doi:10.2147/JMDH.S43952

30. Yusuf S, Collins R, Peto R. Why do we need some large, simple randomized trials? *Stat Med* 1984;3:409-422. doi:10.1002/sim.4780030421
31. Blanco D, Altman D, Moher D, et al. Scoping review on interventions to improve adherence to reporting guidelines in health research. *BMJ Open* 2019;9:e026589. doi:10.1136/bmjopen-2018-026589
32. Barnes C, Boutron I, Giraudeau B, et al. Impact of an online writing aid tool for writing a randomized trial report: the COBWEB (Consort-based WEB tool) randomized controlled trial. *BMC Med* 2015;13:221. doi:10.1186/s12916-015-0460-y
33. Koletsi D, Fleming PS, Behrents RG, et al. The use of tailored subheadings was successful in enhancing compliance with CONSORT in a dental journal. *J Dent* 2017;67:66-71. doi:10.1016/j.jdent.2017.09.009

Figure 1: PRISMA flow diagram

